



Pontryagin Maximum Principle for finite dimensional nonlinear optimal control problems on time scales

Loïc Bourdin, Emmanuel Trélat

► To cite this version:

Loïc Bourdin, Emmanuel Trélat. Pontryagin Maximum Principle for finite dimensional nonlinear optimal control problems on time scales. SIAM Journal on Control and Optimization, 2013, 51 (5), pp.3781–3813. hal-00788477

HAL Id: hal-00788477

<https://hal.science/hal-00788477>

Submitted on 14 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pontryagin Maximum Principle for finite dimensional nonlinear optimal control problems on time scales

Loïc Bourdin*, Emmanuel Trélat†

Abstract

In this article we derive a strong version of the Pontryagin Maximum Principle for general nonlinear optimal control problems on time scales in finite dimension. The final time can be fixed or not, and in the case of general boundary conditions we derive the corresponding transversality conditions. Our proof is based on Ekeland's variational principle. Our statement and comments clearly show the distinction between right-dense points and right-scattered points. At right-dense points a maximization condition of the Hamiltonian is derived, similarly to the continuous-time case. At right-scattered points a weaker condition is derived, in terms of so-called stable Ω -dense directions. We do not make any specific restrictive assumption on the dynamics or on the set Ω of control constraints. Our statement encompasses the classical continuous-time and discrete-time versions of the Pontryagin Maximum Principle, and holds on any general time scale, that is any closed subset of \mathbb{R} .

Keywords: Pontryagin Maximum Principle; optimal control; time scale; transversality conditions; Ekeland's Variational Principle; needle-like variations; right-scattered point; right-dense point.

AMS Classification: 34K35; 34N99; 39A12; 39A13; 49K15; 93C15; 93C55.

1 Introduction

Optimal control theory is concerned with the analysis of controlled dynamical systems, where one aims at steering such a system from a given configuration to some desired target one by minimizing or maximizing some criterion. The Pontryagin Maximum Principle (denoted in short PMP), established at the end of the fifties for finite dimensional general nonlinear continuous-time dynamics (see [44], and see [28] for the history of this discovery), is the milestone of the classical optimal control theory. It provides a first-order necessary condition for optimality, by asserting that any optimal trajectory must be the projection of an extremal. The PMP then reduces the search of optimal trajectories to a boundary value problem posed on extremals. Optimal control theory, and in particular the PMP, have an immense field of applications in various domains, and it is not our aim here to list them. We refer the reader to textbooks on optimal control such as [4, 13, 14, 17, 18, 19, 32, 39, 40, 44, 45, 46, 48] for many examples of theoretical or practical applications of optimal control, essentially in a continuous-time setting.

Right after this discovery the corresponding theory has been developed for discrete-time dynamics, under appropriate convexity assumptions (see e.g. [31, 37, 38]), leading to a version of the

*Laboratoire de Mathématiques et de leurs Applications - Pau (LMAP). UMR CNRS 5142. Université de Pau et des Pays de l'Adour. bourdin.l@univ-pau.fr

†Université Pierre et Marie Curie (Univ. Paris 6) and Institut Universitaire de France, CNRS UMR 7598, Laboratoire Jacques-Louis Lions, F-75005, Paris, France. emmanuel.trelat@upmc.fr

PMP for discrete-time optimal control problems. The considerable development of the discrete-time control theory was motivated by many potential applications e.g. to digital systems or in view of discrete approximations in numerical simulations of differential controlled systems. We refer the reader to the textbooks [12, 22, 43, 46] for details on this theory and many examples of applications. It can be noted that some early works devoted to the discrete-time PMP (like [25]) are mathematically incorrect. Many counter-examples were provided in [12] (see also [43]), showing that, as is now well known, the exact analogous of the continuous-time PMP does not hold at the discrete level. More precisely, the maximization condition of the PMP cannot be expected to hold in general in the discrete-time case. Nevertheless a weaker condition can be derived, see [12, Theorem 42.1 p. 330]. Note as well that approximate maximization conditions are given in [43, Section 6.4] and that a wide literature is devoted to the introduction of convexity assumptions on the dynamics allowing one to recover the maximization condition in the discrete case (such as the concept of *directional convexity* assumption used in [22, 37, 38] for example).

The *time scale* theory was introduced in [33] in order to unify discrete and continuous analysis. A time scale \mathbb{T} is an arbitrary non empty closed subset of \mathbb{R} , and a dynamical system is said to be posed on the time scale \mathbb{T} whenever the time variable evolves along this set \mathbb{T} . The continuous-time case corresponds to $\mathbb{T} = \mathbb{R}$ and the discrete-time case corresponds to $\mathbb{T} = \mathbb{Z}$. The time scale theory aims at closing the gap between continuous and discrete cases and allows one to treat more general models of processes involving both continuous and discrete time elements, and more generally for dynamical systems where the time evolves along a set of a complex nature which may even be a Cantor set (see e.g. [27, 42] for a study of a seasonally breeding population whose generations do not overlap, or [5] for applications to economics). Many notions of standard calculus have been extended to the time scale framework, and we refer the reader to [1, 2, 10, 11] for details on this theory.

The theory of the calculus of variations on time scales, initiated in [8], has been well studied in the existing literature (see e.g. [6, 9, 26, 34, 35]). Few attempts have been made to derive a PMP on time scales. In [36] the authors establish a *weak* PMP for shifted controlled systems, where the controls are not subject to any pointwise constraint and under certain restrictive assumptions. A strong version of the PMP is claimed in [49] but many arguments thereof are erroneous (see Remark 13 for details).

The objective of the present article is to state and prove a strong version of the PMP on time scales, valuable for general nonlinear dynamics, and without assuming any unnecessary Lipschitz or convexity conditions. Our statement is as general as possible, and encompasses the classical continuous-time PMP that can be found e.g. in [40, 44] as well as all versions of discrete-time PMP's mentioned above. In accordance with all known results, the maximization condition is obtained at right-dense points of the time scale and a weaker one (similar to [12, Theorem 42.1 p. 330]) is derived at right-scattered points. Moreover, we consider general constraints on the initial and final values of the state variable and we derive the resulting transversality conditions. We provide as well a version of the PMP for optimal control problems with parameters.

The article is structured as follows. In Section 2, we first provide some basic issues of time scale calculus (Subsection 2.1). We define some appropriate notions such as the notion of stable Ω -dense direction in Subsection 2.2. In Subsection 2.3 we settle the notion of admissible control and define general optimal control problems on time scales. Our main result (Pontryagin Maximum Principle, Theorem 1) is stated in Subsection 2.4, and we analyze and comment the results in a series of remarks. Section 3 is devoted to the proof of Theorem 1. First, in Subsection 3.1 we make some preliminary comments explaining which obstructions may appear when dealing with general time scales, and why we were led to a proof based on Ekeland's Variational Principle. We also comment on the article [49] in Remark 13. In Subsection 3.2, after having shown that the set of admissible controls is open, we define needle-like variations at right-dense and right-scattered

points and derive some properties. In Subsection 3.3, we apply Ekeland's Variational Principle to a well chosen functional in an appropriate complete metric space and then prove the PMP.

2 Main result

Let \mathbb{T} be a time scale, that is, an arbitrary closed subset of \mathbb{R} . We assume throughout that \mathbb{T} is bounded below and that $\text{card}(\mathbb{T}) \geq 2$. We denote by $a = \min \mathbb{T}$.

2.1 Preliminaries on time scales

For every subset A of \mathbb{R} , we denote by $A_{\mathbb{T}} = A \cap \mathbb{T}$. An interval of \mathbb{T} is defined by $I_{\mathbb{T}}$ where I is an interval of \mathbb{R} .

The backward and forward jump operators $\rho, \sigma : \mathbb{T} \rightarrow \mathbb{T}$ are respectively defined by $\rho(t) = \sup\{s \in \mathbb{T} \mid s < t\}$ and $\sigma(t) = \inf\{s \in \mathbb{T} \mid s > t\}$ for every $t \in \mathbb{T}$, where $\rho(a) = a$ and $\sigma(\max \mathbb{T}) = \max \mathbb{T}$ whenever \mathbb{T} is bounded above. A point $t \in \mathbb{T}$ is said to be *left-dense* (respectively, *left-scattered*, *right-dense* or *right-scattered*) whenever $\rho(t) = t$ (respectively, $\rho(t) < t$, $\sigma(t) = t$ or $\sigma(t) > t$). The graininess function $\mu : \mathbb{T} \rightarrow \mathbb{R}^+$ is defined by $\mu(t) = \sigma(t) - t$ for every $t \in \mathbb{T}$. We denote by RS the set of all right-scattered points of \mathbb{T} , and by RD the set of all right-dense points of \mathbb{T} in $\mathbb{T} \setminus \{\sup \mathbb{T}\}$. Note that RS is a subset of $\mathbb{T} \setminus \{\sup \mathbb{T}\}$, is at most countable (see [21, Lemma 3.1]), and note that RD is the complement of RS in $\mathbb{T} \setminus \{\sup \mathbb{T}\}$. For every $b \in \mathbb{T} \setminus \{a\}$ and every $s \in [a, b]_{\mathbb{T}} \cap \text{RD}$, we set

$$\mathcal{V}_s^b = \{\beta \geq 0, s + \beta \in [s, b]_{\mathbb{T}}\}. \quad (1)$$

Note that 0 is not isolated in \mathcal{V}_s^b .

Δ -differentiability. We set $\mathbb{T}^\kappa = \mathbb{T} \setminus \{\max \mathbb{T}\}$ whenever \mathbb{T} admits a left-scattered maximum, and $\mathbb{T}^\kappa = \mathbb{T}$ otherwise. Let $n \in \mathbb{N}^*$; a function $q : \mathbb{T} \rightarrow \mathbb{R}^n$ is said to be Δ -differentiable at $t \in \mathbb{T}^\kappa$ if the limit

$$q^\Delta(t) = \lim_{\substack{s \rightarrow t \\ s \in \mathbb{T}}} \frac{q^\sigma(t) - q(s)}{\sigma(t) - s}$$

exists in \mathbb{R}^n , where $q^\sigma = q \circ \sigma$. Recall that, if $t \in \mathbb{T}^\kappa$ is a right-dense point of \mathbb{T} , then q is Δ -differentiable at t if and only if the limit of $\frac{q(t) - q(s)}{t - s}$ as $s \rightarrow t$, $s \in \mathbb{T}$, exists; in that case it is equal to $q^\Delta(t)$. If $t \in \mathbb{T}^\kappa$ is a right-scattered point of \mathbb{T} and if q is continuous at t , then q is Δ -differentiable at t , and $q^\Delta(t) = (q^\sigma(t) - q(t))/\mu(t)$ (see [10]).

If $q, q' : \mathbb{T} \rightarrow \mathbb{R}^n$ are both Δ -differentiable at $t \in \mathbb{T}^\kappa$, then the scalar product $\langle q, q' \rangle_{\mathbb{R}^n}$ is Δ -differentiable at t and

$$\langle q, q' \rangle_{\mathbb{R}^n}^\Delta(t) = \langle q^\Delta(t), q'^\sigma(t) \rangle_{\mathbb{R}^n} + \langle q(t), q'^\Delta(t) \rangle_{\mathbb{R}^n} = \langle q^\Delta(t), q'(t) \rangle_{\mathbb{R}^n} + \langle q^\sigma(t), q'^\Delta(t) \rangle_{\mathbb{R}^n} \quad (2)$$

(Leibniz formula, see [10, Theorem 1.20]).

Lebesgue Δ -measure and Lebesgue Δ -integrability. Let μ_Δ be the Lebesgue Δ -measure on \mathbb{T} defined in terms of Carathéodory extension in [11, Chapter 5]. We also refer the reader to [3, 21, 29] for more details on the μ_Δ -measure theory. For all $(c, d) \in \mathbb{T}^2$ such that $c \leq d$, there holds $\mu_\Delta([c, d]_{\mathbb{T}}) = d - c$. Recall that $A \subset \mathbb{T}$ is a μ_Δ -measurable set of \mathbb{T} if and only if A is an usual μ_L -measurable set of \mathbb{R} , where μ_L denotes the usual Lebesgue measure (see [21, Proposition 3.1]). Moreover, if $A \subset \mathbb{T} \setminus \{\sup \mathbb{T}\}$, then

$$\mu_\Delta(A) = \mu_L(A) + \sum_{r \in A \cap \text{RS}} \mu(r).$$

Let $A \subset \mathbb{T}$. A property is said to hold Δ -almost everywhere (in short, Δ -a.e.) on A if it holds for every $t \in A \setminus A'$, where $A' \subset A$ is some μ_Δ -measurable subset of \mathbb{T} satisfying $\mu_\Delta(A') = 0$. In particular, since $\mu_\Delta(\{r\}) = \mu(r) > 0$ for every $r \in \text{RS}$, we conclude that if a property holds Δ -a.e. on A , then it holds for every $r \in A \cap \text{RS}$.

Let $n \in \mathbb{N}^*$ and let $A \subset \mathbb{T} \setminus \{\sup \mathbb{T}\}$ be a μ_Δ -measurable subset of \mathbb{T} . Consider a function q defined Δ -a.e. on A with values in \mathbb{R}^n . Let $A_0 = A \cup]r, \sigma(r)[_{r \in A \cap \text{RS}}$, and let q_0 be the extension of q defined μ_L -a.e. on A_0 by $q_0(t) = q(t)$ whenever $t \in A$, and by $q_0(t) = q(r)$ whenever $t \in]r, \sigma(r)[$, for every $r \in A \cap \text{RS}$. Recall that q is μ_Δ -measurable on A if and only if q_0 is μ_L -measurable on A_0 (see [21, Proposition 4.1]).

The functional space $L_\mathbb{T}^\infty(A, \mathbb{R}^n)$ is the set of all functions q defined Δ -a.e. on A , with values in \mathbb{R}^n , that are μ_Δ -measurable on A and bounded almost everywhere. Endowed with the norm $\|q\|_{L_\mathbb{T}^\infty(A, \mathbb{R}^n)} = \sup_{\tau \in A} \|q(\tau)\|_{\mathbb{R}^n}$, it is a Banach space (see [3, Theorem 2.5]). Here the notation $\|\cdot\|_{\mathbb{R}^n}$ stands for the usual Euclidean norm of \mathbb{R}^n . The functional space $L_\mathbb{T}^1(A, \mathbb{R}^n)$ is the set of all functions q defined Δ -a.e. on A , with values in \mathbb{R}^n , that are μ_Δ -measurable on A and such that $\int_A \|q(\tau)\|_{\mathbb{R}^n} \Delta\tau < +\infty$. Endowed with the norm $\|q\|_{L_\mathbb{T}^1(A, \mathbb{R}^n)} = \int_A \|q(\tau)\|_{\mathbb{R}^n} \Delta\tau$, it is a Banach space (see [3, Theorem 2.5]). We recall here that if $q \in L_\mathbb{T}^1(A, \mathbb{R}^n)$ then

$$\int_A q(\tau) \Delta\tau = \int_{A_0} q_0(\tau) d\tau = \int_A q(\tau) d\tau + \sum_{r \in A \cap \text{RS}} \mu(r) q(r)$$

(see [21, Theorems 5.1 and 5.2]). Note that if A is bounded then $L_\mathbb{T}^\infty(A, \mathbb{R}^n) \subset L_\mathbb{T}^1(A, \mathbb{R}^n)$. The functional space $L_{\text{loc}, \mathbb{T}}^\infty(\mathbb{T} \setminus \{\sup \mathbb{T}\}, \mathbb{R}^n)$ is the set of all functions q defined Δ -a.e. on $\mathbb{T} \setminus \{\sup \mathbb{T}\}$, with values in \mathbb{R}^n , that are μ_Δ -measurable on $\mathbb{T} \setminus \{\sup \mathbb{T}\}$ and such that $q \in L_\mathbb{T}^\infty([c, d[_\mathbb{T}, \mathbb{R}^n)$ for all $(c, d) \in \mathbb{T}^2$ such that $c < d$.

Absolutely continuous functions. Let $n \in \mathbb{N}^*$ and let $(c, d) \in \mathbb{T}^2$ such that $c < d$. Let $\mathcal{C}([c, d]_\mathbb{T}, \mathbb{R}^n)$ denote the space of continuous functions defined on $[c, d]_\mathbb{T}$ with values in \mathbb{R}^n . Endowed with its usual uniform norm $\|\cdot\|_\infty$, it is a Banach space. Let $\text{AC}([c, d]_\mathbb{T}, \mathbb{R}^n)$ denote the subspace of absolutely continuous functions.

Let $t_0 \in [c, d]_\mathbb{T}$ and $q : [c, d]_\mathbb{T} \rightarrow \mathbb{R}^n$. It is easily derived from [20, Theorem 4.1] that $q \in \text{AC}([c, d]_\mathbb{T}, \mathbb{R}^n)$ if and only if q is Δ -differentiable Δ -a.e. on $[c, d[_\mathbb{T}$ and satisfies $q^\Delta \in L_\mathbb{T}^1([c, d[_\mathbb{T}, \mathbb{R}^n)$, and for every $t \in [c, d]_\mathbb{T}$ there holds $q(t) = q(t_0) + \int_{[t_0, t[_\mathbb{T}} q^\Delta(\tau) \Delta\tau$ whenever $t \geq t_0$, and $q(t) = q(t_0) - \int_{[t, t_0[_\mathbb{T}} q^\Delta(\tau) \Delta\tau$ whenever $t \leq t_0$.

Assume that $q \in L_\mathbb{T}^1([c, d]_\mathbb{T}, \mathbb{R}^n)$, and let Q be the function defined on $[c, d]_\mathbb{T}$ by $Q(t) = \int_{[t_0, t[_\mathbb{T}} q(\tau) \Delta\tau$ whenever $t \geq t_0$, and by $Q(t) = -\int_{[t, t_0[_\mathbb{T}} q(\tau) \Delta\tau$ whenever $t \leq t_0$. Then $Q \in \text{AC}([c, d]_\mathbb{T})$ and $Q^\Delta = q$ Δ -a.e. on $[c, d[_\mathbb{T}$.

Note that, if $q \in \text{AC}([c, d]_\mathbb{T}, \mathbb{R}^n)$ is such that $q^\Delta = 0$ Δ -a.e. on $[c, d[_\mathbb{T}$, then q is constant on $[c, d]_\mathbb{T}$, and that, if $q, q' \in \text{AC}([c, d]_\mathbb{T}, \mathbb{R}^n)$, then $\langle q, q' \rangle_{\mathbb{R}^n} \in \text{AC}([c, d]_\mathbb{T}, \mathbb{R})$ and the Leibniz formula (2) is available Δ -a.e. on $[c, d[_\mathbb{T}$.

For every $q \in L_\mathbb{T}^1([c, d]_\mathbb{T}, \mathbb{R}^n)$, let $\mathcal{L}_{[c, d]_\mathbb{T}}(q)$ be the set of points $t \in [c, d]_\mathbb{T}$ that are Δ -Lebesgue points of q . There holds $\mu_\Delta(\mathcal{L}_{[c, d]_\mathbb{T}}(q)) = \mu_\Delta([c, d]_\mathbb{T}) = d - c$, and

$$\lim_{\substack{\beta \rightarrow 0^+ \\ \beta \in \mathcal{V}_s^d}} \frac{1}{\beta} \int_{[s, s+\beta[_\mathbb{T}} q(\tau) \Delta\tau = q(s), \quad (3)$$

for every $s \in \mathcal{L}_{[c, d]_\mathbb{T}}(q) \cap \text{RD}$, where \mathcal{V}_s^d is defined by (1).

Remark 1. Note that the analogous result for $s \in \mathcal{L}_{[c, d]_\mathbb{T}}(q) \cap \text{LD}$ is not true in general. Indeed, let $q \in L_\mathbb{T}^1([c, d]_\mathbb{T}, \mathbb{R}^n)$ and assume that there exists a point $s \in [c, d]_\mathbb{T} \cap \text{LD} \cap \text{RS}$. Since $\mu_\Delta(\{s\}) =$

$\mu(s) > 0$, one has $s \in \mathcal{L}_{[c,d[\mathbb{T}}(q)$. Nevertheless the limit $\frac{1}{\beta} \int_{[s-\beta, s[\mathbb{T}} q(\tau) \Delta\tau$ as $\beta \rightarrow 0^+$ with $s-\beta \in \mathbb{T}$, is not necessarily equal to $q(s)$. For instance, consider $\mathbb{T} = [0, 1] \cup \{2\}$, $s = 1$ and q defined on \mathbb{T} by $q(t) = 0$ for every $t \neq 1$ and $q(1) = 1$.

Remark 2. Recall that two distinct derivative operators are usually considered in the time scale calculus, namely, the Δ -derivative, corresponding to a forward derivative, and the ∇ -derivative, corresponding to a backward derivative, and that both of them are associated with a notion of integral. In this article, without loss of generality we consider optimal control problems defined on time scales with a Δ -derivative and with a cost function written with the corresponding notion of Δ -integral. Our main result, the PMP, is then stated using the notions of right-dense and right-scattered points. All problems and results of our article can be as well stated in terms of ∇ -derivative, ∇ -integral, left-dense and left-scattered points.

2.2 Topological preliminaries

Let $m \in \mathbb{N}^*$ and let Ω be a non empty closed subset of \mathbb{R}^m . In this section we define the notion of stable Ω -dense direction. In our main result the set Ω will stand for the set of pointwise constraints on the controls.

Definition 1. Let $v \in \Omega$ and $v' \in \mathbb{R}^m$.

1. We set $\mathcal{D}^\Omega(v, v') = \{0 \leq \alpha \leq 1 \mid v + \alpha(v' - v) \in \Omega\}$. Note that $0 \in \mathcal{D}^\Omega(v, v')$.
2. We say that v' is a Ω -dense direction from v if 0 is not isolated in $\mathcal{D}^\Omega(v, v')$. The set of all Ω -dense directions from v is denoted by $\mathcal{D}^\Omega(v)$.
3. We say that v' is a stable Ω -dense direction from v if there exists $\varepsilon > 0$ such that $v' \in \mathcal{D}^\Omega(v'')$ for every $v'' \in \overline{B}(v, \varepsilon) \cap \Omega$, where $\overline{B}(v, \varepsilon)$ is the closed ball of \mathbb{R}^m centered at v and with radius ε . The set of all stable Ω -dense directions from v is denoted by $\mathcal{D}_{\text{stab}}^\Omega(v)$.

Note that $v' \in \mathcal{D}_{\text{stab}}^\Omega(v)$ means that v' is a Ω -dense direction from v'' for every $v'' \in \Omega$ in a neighbourhood of v . In the following, we denote by Int the interior of a subset. We have the following easy properties.

1. If $v \in \text{Int}(\Omega)$, then $\mathcal{D}_{\text{stab}}^\Omega(v) = \mathbb{R}^m$.
2. If $\Omega = \{v\}$ then $\mathcal{D}_{\text{stab}}^\Omega(v) = \{v\}$;
3. If Ω is convex then $\Omega \subset \mathcal{D}_{\text{stab}}^\Omega(v)$ for every $v \in \Omega$.

For every $v \in \Omega$, we denote by $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v))$ the closed convex cone of vertex v spanned by $\mathcal{D}_{\text{stab}}^\Omega(v)$, with the agreement that $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \{v\}$ whenever $\mathcal{D}_{\text{stab}}^\Omega(v) = \emptyset$. In particular, there holds $v \in \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v))$ for every $v \in \Omega$.

Although elementary, since these notions are new (up to our knowledge), before proceeding with our main result (stated in Section 2.3) we provide the reader with several simple examples illustrating these notions. Since $\mathcal{D}_{\text{stab}}^\Omega(v) = \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \mathbb{R}^m$ for every $v \in \text{Int}(\Omega)$, we focus on elements $v \in \partial\Omega$ in the examples below.

Example 1. Assume that $m = 1$. The closed convex subsets Ω of \mathbb{R} having a nonempty interior and such that $\partial\Omega \neq \emptyset$ are closed intervals bounded above or below and not reduced to a singleton. If Ω is bounded below then $\mathcal{D}_{\text{stab}}^\Omega(\min \Omega) = \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(\min \Omega)) = [\min \Omega, +\infty[$, and if Ω is bounded above then $\mathcal{D}_{\text{stab}}^\Omega(\max \Omega) = \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(\max \Omega)) =]-\infty, \max \Omega]$.

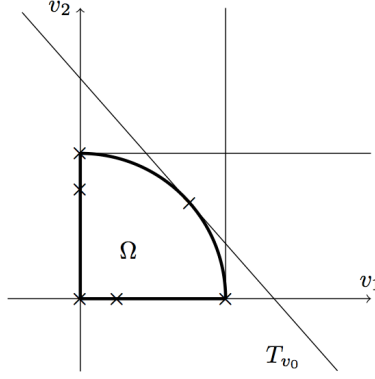


Figure 1:

Example 2. Assume that $m = 2$ and let Ω be the convex set of $v = (v_1, v_2) \in \mathbb{R}^2$ such that $v_1 \geq 0$, $v_2 \geq 0$ and $v_1^2 + v_2^2 \leq 1$ (see Figure 1). The stable Ω -dense directions for elements $v \in \partial\Omega$ are given by:

- if $v = (0, 0)$, then $\mathcal{D}_{\text{stab}}^\Omega(v) = \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = (\mathbb{R}^+)^2$;
- if $v = (0, v_0)$ with $0 < v_0 < 1$, then $\mathcal{D}_{\text{stab}}^\Omega(v) = \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \mathbb{R}^+ \times \mathbb{R}$;
- if $v = (v_0, 0)$ with $0 < v_0 < 1$, then $\mathcal{D}_{\text{stab}}^\Omega(v) = \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \mathbb{R} \times \mathbb{R}^+$;
- if $v = (0, 1)$, then $\mathcal{D}_{\text{stab}}^\Omega(v) = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1 \geq 0, v_2 < 1\} \cup \{v\}$ and $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1 \geq 0, v_2 \leq 1\}$;
- if $v = (1, 0)$, then $\mathcal{D}_{\text{stab}}^\Omega(v) = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1 < 1, v_2 \geq 0\} \cup \{v\}$ and $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1 \leq 1, v_2 \geq 0\}$;
- if $v = (v_0, \sqrt{1 - v_0^2})$ with $0 < v_0 < 1$, then $\mathcal{D}_{\text{stab}}^\Omega(v)$ is the union of $\{v\}$ and of the strict hypograph of T_{v_0} , and $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v))$ is the hypograph of T_{v_0} .

Remark 3. Let Ω be a non empty closed convex subset of \mathbb{R}^m and let $\text{Aff}(\Omega)$ denote the smallest affine subspace of \mathbb{R}^m containing Ω . For every $v \in \partial\Omega$ that is not a corner point, $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v))$ is the half-space of $\text{Aff}(\Omega)$ delimited by the tangent hyperplane (in $\text{Aff}(\Omega)$) of Ω at v , and containing Ω .

Example 3. Assume that $m = 2$ and let Ω be the set of $v = (v_1, v_2) \in \mathbb{R}^2$ such that $v_2 \leq |v_1|$ (see Figure 2). The stable Ω -dense directions for elements $v \in \partial\Omega$ are given by:

- if $v = (v_0, |v_0|)$ with $v_0 < 0$, then $\mathcal{D}^\Omega(v) = \mathcal{D}_{\text{stab}}^\Omega(v) = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_2 \leq -v_1\}$;
- if $v = (v_0, |v_0|)$ with $v_0 > 0$, then $\mathcal{D}^\Omega(v) = \mathcal{D}_{\text{stab}}^\Omega(v) = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_2 \leq v_1\}$;
- if $v = (0, 0)$, then $\mathcal{D}^\Omega(v) = \Omega$, $\mathcal{D}_{\text{stab}}^\Omega(v) = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_2 \leq -|v_1|\}$;

Note that, in all cases, $\mathcal{D}_{\text{stab}}^\Omega(v)$ is a closed convex cone of vertex v and therefore $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \mathcal{D}_{\text{stab}}^\Omega(v)$.

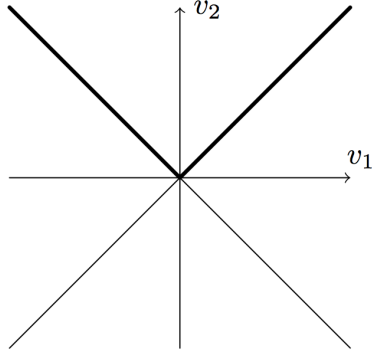


Figure 2:

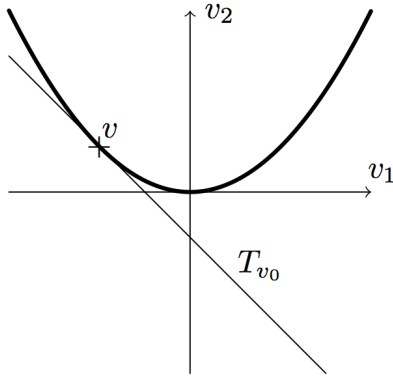


Figure 3:

Example 4. Assume that $m = 2$ and let Ω be the set of $v = (v_1, v_2) \in \mathbb{R}^2$ such that $v_2 \leq v_1^2$ (see Figure 3). Let $v_0 \in \mathbb{R}$ and let $T_{v_0}(v_1) = v_0(2v_1 - v_0)$ denote the graph of the tangent to Ω at the point $v = (v_0, v_0^2)$. It is easy to see that $\mathcal{D}^\Omega(v)$ is the hypograph of T_{v_0} , that $\mathcal{D}_{\text{stab}}^\Omega(v)$ is the strict hypograph of T_{v_0} (note that $v \notin \mathcal{D}_{\text{stab}}^\Omega(v)$), and that $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v))$ is the hypograph of T_{v_0} .

Remark 4. The above example shows that it may happen that $v \notin \mathcal{D}_{\text{stab}}^\Omega(v)$. Actually, it may happen that $\mathcal{D}_{\text{stab}}^\Omega(v) = \emptyset$. For example, if Ω is the unit sphere of \mathbb{R}^2 , then $\mathcal{D}_{\text{stab}}^\Omega(v) = \emptyset$ for every $v \in \Omega$, and hence $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \{v\}$.

Example 5. Assume that $m = 2$. We set $\Omega = \cup_{k \in \mathbb{N}} \overline{\Omega}_k \cup \overline{\Omega}_\infty$, where $\Omega_k = \{(v_1, (1 - v_1)/2^k) \mid 0 < v_1 < 1\}$ for every $k \in \mathbb{N}$, and $\Omega_\infty = \{(v_1, 0) \mid 0 < v_1 < 1\}$ (see Figure 4). Note that Ω has an empty interior. Denote by $\bar{v} = (1, 0)$. We have the following properties:

- if $v \in \Omega_k$ with $k \in \mathbb{N}$, then $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \mathcal{D}_{\text{stab}}^\Omega(v) = \mathcal{D}^\Omega(v) = \{(v_1, (1 - v_1)/2^k) \mid v_1 \in \mathbb{R}\}$;
- if $v = (0, 1/2^k)$ with $k \in \mathbb{N}$, then $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \mathcal{D}_{\text{stab}}^\Omega(v) = \mathcal{D}^\Omega(v) = \{(v_1, (1 - v_1)/2^k) \mid v_1 \geq 0\}$;

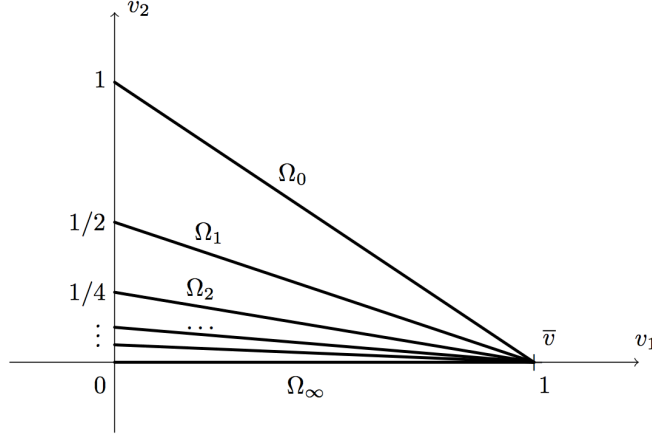


Figure 4:

- if $v = (v_1, 0)$ with $0 < v_1 < 1$, then $\mathcal{D}^\Omega(v) = \mathbb{R} \times \mathbb{R}^+$ and $\mathcal{D}_{\text{stab}}^\Omega(v) = \{\bar{v}\}$, and thus $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = [v_1, +\infty[\times \{0\}$;
- if $v = (0, 0)$, then $\mathcal{D}^\Omega(v) = (\mathbb{R}^+)^2$ and $\mathcal{D}_{\text{stab}}^\Omega(v) = \{\bar{v}\}$, and thus $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \mathbb{R}^+ \times \{0\}$;
- if $v = \bar{v}$, then $\mathcal{D}^\Omega(\bar{v}) = \cup_{k \in \mathbb{N}} \{(v_1, (1-v_1)/2^k) \mid v_1 \leq 1\} \cup \{(v_1, 0) \mid v_1 \leq 1\}$ and $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v)) = \mathcal{D}_{\text{stab}}^\Omega(\bar{v}) = \{\bar{v}\}$.

2.3 General nonlinear optimal control problem on time scales

Let n and m be nonzero integers, and let Ω be a non empty closed subset of \mathbb{R}^m . Throughout the article, we consider the general nonlinear control system on the time scale \mathbb{T}

$$q^\Delta(t) = f(q(t), u(t), t), \quad (4)$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{T} \rightarrow \mathbb{R}^n$ is a continuous function of class \mathcal{C}^1 with respect to its two first variables, and where the control functions u belong to $L_{\text{loc}, \mathbb{T}}^\infty(\mathbb{T} \setminus \{\sup \mathbb{T}\}; \Omega)$.

Before defining an optimal control problem associated with the control system (4), the first question that has to be addressed is the question of the existence and uniqueness of a solution of (4), for a given control function and a given initial condition $q(a) = q_a \in \mathbb{R}^n$. Since there did not exist up to now in the existing literature any Cauchy-Lipschitz like theorem, sufficiently general to cover such a situation, in the companion paper [16] we derived a general Cauchy-Lipschitz (or Picard-Lindelöf) theorem for general nonlinear systems posed on time scales, providing existence and uniqueness of the maximal solution of a given Δ -Cauchy problem under suitable assumptions like regressivity and local Lipschitz continuity, and discussed some related issues like the behavior of maximal solutions at terminal points.

Setting $\mathcal{U} = L_{\text{loc}, \mathbb{T}}^\infty(\mathbb{T} \setminus \{\sup \mathbb{T}\}; \mathbb{R}^m)$, let us first recall the notion of a solution of (4), for a given control $u \in \mathcal{U}$ (see [16, Definitions 6 and 7]). The couple $(q, I_{\mathbb{T}})$ is said to be a solution of (4) if $I_{\mathbb{T}}$ is an interval of \mathbb{T} satisfying $a \in I_{\mathbb{T}}$ and $I_{\mathbb{T}} \setminus \{a\} \neq \emptyset$, if $q \in \text{AC}([a, b]_{\mathbb{T}}, \mathbb{R}^n)$ and (4) holds for Δ -a.e. $t \in [a, b]_{\mathbb{T}}$, for every $b \in I_{\mathbb{T}} \setminus \{a\}$.

According to [16, Theorem 1], for every control $u \in \mathcal{U}$ and every $q_a \in \mathbb{R}^n$, there exists a unique maximal solution $q(\cdot, u, q_a)$ of (4), such that $q(a) = q_a$, defined on the maximal interval $I_{\mathbb{T}}(u, q_a)$. The word *maximal* means that $q(\cdot, u, q_a)$ is an extension of any other solution. Note

that $q(t, u, q_a) = q_a + \int_{[a, t]_{\mathbb{T}}} f(q(\tau, u, q_a), u(\tau), \tau) \Delta\tau$, for every $t \in I_{\mathbb{T}}(u, q_a)$ (see [16, Lemma 1]), and that either $I_{\mathbb{T}}(u, q_a) = \mathbb{T}$, that is, $q(\cdot, u, q_a)$ is a *global* solution of (4), or $I_{\mathbb{T}}(u, q_a) = [a, b]_{\mathbb{T}}$ where $b \in \mathbb{T} \setminus \{a\}$ is a left-dense point of \mathbb{T} , and in this case, $q(\cdot, u, q_a)$ is not bounded on $I_{\mathbb{T}}(u, q_a)$ (see [16, Theorem 2]).

These results are instrumental to define the concept of admissible control.

Definition 2. For every $q_a \in \mathbb{R}^n$, the control $u \in \mathcal{U}$ is said to be *admissible* on $[a, b]_{\mathbb{T}}$ for some given $b \in \mathbb{T} \setminus \{a\}$ whenever $q(\cdot, u, q_a)$ is well defined on $[a, b]_{\mathbb{T}}$, that is, $b \in I_{\mathbb{T}}(u, q_a)$.

We are now in a position to define rigorously a general optimal control problem on the time scale \mathbb{T} .

Let $j \in \mathbb{N}^*$ and S be a non empty closed convex subset of \mathbb{R}^j . Let $f^0 : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{T} \rightarrow \mathbb{R}$ be a continuous function of class \mathcal{C}^1 with respect to its two first variables, and $g : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^j$ be a function of class \mathcal{C}^1 . In what follows the subset S and the function g account for constraints on the initial and final conditions of the control problem.

Throughout the article, we consider the optimal control problem on \mathbb{T} , denoted in short $(\mathbf{OCP})_{\mathbb{T}}$, of determining a trajectory $q^*(\cdot)$ defined on $[a, b^*]_{\mathbb{T}}$, solution of (4) and associated with a control $u^* \in L_{\mathbb{T}}^{\infty}([a, b^*]_{\mathbb{T}}; \Omega)$, minimizing the cost function

$$C(b, u) = \int_{[a, b]_{\mathbb{T}}} f^0(q(\tau), u(\tau), \tau) \Delta\tau \quad (5)$$

over all possible trajectories $q(\cdot)$ defined on $[a, b]_{\mathbb{T}}$, solutions of (4) and associated with an admissible control $u \in L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}; \Omega)$, with $b \in \mathbb{T} \setminus \{a\}$, and satisfying $g(q(a), q(b)) \in S$.

The final time can be fixed or not. If it is fixed then $b^* = b$ in $(\mathbf{OCP})_{\mathbb{T}}$.

2.4 Pontryagin Maximum Principle

In the statement below, the *orthogonal* of S at a point $x \in S$ is defined by

$$\mathcal{O}_S(x) = \{x' \in \mathbb{R}^j \mid \forall x'' \in S, \langle x', x'' - x \rangle_{\mathbb{R}^j} \leq 0\}. \quad (6)$$

It is a closed convex cone containing 0.

The *Hamiltonian* of the optimal control problem $(\mathbf{OCP})_{\mathbb{T}}$ is the function $H : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$ defined by $H(q, u, p, p^0, t) = \langle p, f(q, u, t) \rangle_{\mathbb{R}^n} + p^0 f^0(q, u, t)$.

Theorem 1 (Pontryagin Maximum Principle). *Let $b^* \in \mathbb{T} \setminus \{a\}$. If the trajectory $q^*(\cdot)$, defined on $[a, b^*]_{\mathbb{T}}$ and associated with a control $u^* \in L_{\mathbb{T}}^{\infty}([a, b^*]_{\mathbb{T}}; \Omega)$, is a solution of $(\mathbf{OCP})_{\mathbb{T}}$, then there exist $p^0 \leq 0$ and $\psi \in \mathbb{R}^j$, with $(p^0, \psi) \neq (0, 0)$, and there exists a mapping $p(\cdot) \in AC([a, b^*]_{\mathbb{T}}, \mathbb{R}^n)$ (called adjoint vector), such that there holds*

$$q^{*\Delta}(t) = \frac{\partial H}{\partial p}(q^*(t), u^*(t), p^{\sigma}(t), p^0, t), \quad p^{\Delta}(t) = -\frac{\partial H}{\partial q}(q^*(t), u^*(t), p^{\sigma}(t), p^0, t), \quad (7)$$

for Δ -a.e. $t \in [a, b^*]_{\mathbb{T}}$. Moreover, there holds

$$\left\langle \frac{\partial H}{\partial u}(q^*(r), u^*(r), p^{\sigma}(r), p^0, r), v - u^*(r) \right\rangle_{\mathbb{R}^m} \leq 0, \quad (8)$$

for every $r \in [a, b^*]_{\mathbb{T} \cap \mathbb{R}S}$ and every $v \in \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^{\Omega}(u^*(r)))$, and

$$H(q^*(s), u^*(s), p^{\sigma}(s), p^0, s) = \max_{v \in \Omega} H(q^*(s), v, p^{\sigma}(s), p^0, s), \quad (9)$$

for Δ -a.e. $s \in [a, b^*[_\mathbb{T} \cap \text{RD}$.

Besides, one has the transversality conditions on the initial and final adjoint vector

$$p(a) = -\left(\frac{\partial g}{\partial q_1}(q^*(a), q^*(b^*))\right)^\top \psi, \quad p(b^*) = \left(\frac{\partial g}{\partial q_2}(q^*(a), q^*(b^*))\right)^\top \psi, \quad (10)$$

and $-\psi \in \mathcal{O}_S(g(q^*(a), q^*(b^*)))$.

Furthermore, if the final time b^* is not fixed in $(\mathbf{OCP})_\mathbb{T}$, and if additionally b^* belongs to the interior of \mathbb{T} for the topology of \mathbb{R} , then

$$\max_{v \in \Omega} H(q^*(b^*), v, p^\sigma(b^*), p^0, b^*) = 0, \quad (11)$$

and if H is moreover autonomous (that is, does not depend on t), then

$$\int_{[a, b^*[_\mathbb{T}} H(q^*(t), u^*(t), p^\sigma(t), p^0) \Delta t = 0. \quad (12)$$

Theorem 1 is proved in Section 3.

Remark 5 (PMP for optimal control problems with parameters). Before proceeding with a series of remarks and comments, we provide a version of the PMP for optimal control problems with parameters. Let Λ be a Banach space. We consider the general nonlinear control system with parameters on the time scale \mathbb{T}

$$q^\Delta(t) = f(\lambda, q(t), u(t), t), \quad (13)$$

where $f : \Lambda \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{T} \rightarrow \mathbb{R}^n$ is a continuous function of class \mathcal{C}^1 with respect to its three first variables, and where $u \in \mathcal{U}$ as before. The notion of admissibility is defined as before. Let $f^0 : \Lambda \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{T} \rightarrow \mathbb{R}$ be a continuous function of class \mathcal{C}^1 with respect to its three first variables, and $g : \Lambda \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^j$ be a function of class \mathcal{C}^1 .

We consider the optimal control problem on \mathbb{T} , denoted in short $(\mathbf{OCP})_\mathbb{T}^\lambda$, of determining a trajectory $q^*(\cdot)$ defined on $[a, b^*]_\mathbb{T}$, solution of (13) and associated with a control $u^* \in L_\mathbb{T}^\infty([a, b^*]_\mathbb{T}; \Omega)$ and with a parameter $\lambda^* \in \Lambda$, minimizing the cost function $C(\lambda, b, u) = \int_{[a, b]_\mathbb{T}} f^0(\lambda, q(\tau), u(\tau), \tau) \Delta \tau$ over all possible trajectories $q(\cdot)$ defined on $[a, b]_\mathbb{T}$, solutions of (13) and associated with $\lambda \in \Lambda$ and with an admissible control $u \in L_\mathbb{T}^\infty([a, b]_\mathbb{T}; \Omega)$, with $b \in \mathbb{T} \setminus \{a\}$, and satisfying $g(\lambda, q(a), q(b)) \in S$. The final time can be fixed or not.

The *Hamiltonian* of $(\mathbf{OCP})_\mathbb{T}^\lambda$ is the function $H : \Lambda \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}$ defined by

$$H(\lambda, q, u, p, p^0, t) = \langle p, f(\lambda, q, u, t) \rangle_{\mathbb{R}^n} + p^0 f^0(\lambda, q, u, t).$$

If the trajectory $q^*(\cdot)$, defined on $[a, b^*]_\mathbb{T}$ and associated with a control $u^* \in L_\mathbb{T}^\infty([a, b^*]_\mathbb{T}; \Omega)$ and with a parameter $\lambda^* \in \Lambda$, is a solution of $(\mathbf{OCP})_\mathbb{T}^\lambda$, then all conclusions of Theorem 1 (except (12)) hold, and moreover

$$\int_{[a, b^*[_\mathbb{T}} \frac{\partial H}{\partial \lambda}(\lambda^*, q^*(t), u^*(t), p^\sigma(t), p^0, t) \Delta t + \left\langle \frac{\partial g}{\partial \lambda}(\lambda^*, q^*(a), q^*(b^*)), \psi \right\rangle_{\mathbb{R}^j} = 0. \quad (14)$$

This additional statement is proved as well in Section 3.

Remark 6. As is well known, the Lagrange multiplier (p^0, ψ) (and thus the triple $(p(\cdot), p^0, \psi)$) is defined up to a multiplicative scalar. Defining as usual an *extremal* as a quadruple $(q(\cdot), u(\cdot), p(\cdot), p^0)$ solution of the above equations, an extremal is said to be *normal* whenever $p^0 \neq 0$ and *abnormal* whenever $p^0 = 0$. The component p^0 corresponds to the Lagrange multiplier associated with the cost function. In the normal case $p^0 \neq 0$ it is usual to normalize the Lagrange multiplier so that $p^0 = -1$. Finally, note that the convention $p^0 \leq 0$ in the PMP leads to a maximization condition of the Hamiltonian (the convention $p^0 \geq 0$ would lead to a minimization condition).

Remark 7. As already mentioned in Remark 2, without loss of generality we consider in this article optimal control problems defined with the notion of Δ -derivative and Δ -integral. These notions are naturally associated with the concepts of right-dense and right-scattered points in the basic properties of calculus (see Section 2.1). Therefore, when using a Δ -derivative in the definition of $(\text{OCP})_{\mathbb{T}}$ one cannot hope to derive in general, for instance, a maximization condition at left-dense points (see the counterexample of Remark 1).

Remark 8. In the classical continuous-time setting, it is well known that the maximized Hamiltonian along the optimal extremal, that is, the function $t \mapsto \max_{v \in \Omega} H(q^*(t), v, p^\sigma(t), p^0, t)$, is Lipschitzian on $[a, b^*]$, and if the dynamics are autonomous (that is, if H does not depend on t) then this function is constant. Moreover, if the final time is free then the maximized Hamiltonian vanishes at the final time.

In the discrete-time setting and a fortiori in the general time scale setting, none of these properties do hold any more in general (see Examples 6 and 8 below). The non constant feature is due in particular to the fact that the usual formula of derivative of a composition does not hold in general time scale calculus.

Remark 9. The PMP is derived here in a general framework. We do not make any particular assumption on the time scale \mathbb{T} , and do not assume that the set of control constraints Ω is convex or compact. In Section 3.1, we discuss the strategy of proof of Theorem 1 and we explain how the generality of the framework led us to choose a method based on a variational principle rather than one based on a fixed-point theorem.

We do not make any convexity assumption on the dynamics (f, f^0) . As a consequence, and as is well known in the discrete case (see e.g. [12, p. 50–63]), at right-scattered points the maximization condition (9) does not hold true in general and must be weakened into (8) (see Remark 11).

Remark 10. The inequality (8), valuable at right-scattered points, can be written as

$$\frac{\partial H}{\partial u}(q^*(r), u^*(r), p^\sigma(r), p^0, r) \in \mathcal{O}_{\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(u^*(r)))}(u^*(r)).$$

In particular, if $u^*(r) \in \text{Int}(\Omega)$ then $\frac{\partial H}{\partial u}(q^*(r), u^*(r), p^\sigma(r), p^0, r) = 0$. This equality holds true at every right-scattered point if for instance $\Omega = \mathbb{R}^m$ (and also at right-dense points: this is the context of what is usually referred to as the *weak PMP*, see [36] where this weaker result is derived on general time scales for shifted control systems).

If Ω is convex, since $u^*(r) \in \Omega \subset \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(u^*(r)))$, then there holds in particular

$$\frac{\partial H}{\partial u}(q^*(r), u^*(r), p^\sigma(r), p^0, r) \in \mathcal{O}_\Omega(u^*(r)),$$

for every $r \in [a, b]_{\mathbb{T} \cap \text{RS}}$.

Note that, if the inequality (8) is strict then $u^*(r)$ satisfies a local maximization condition on $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(u^*(r)))$ (see also [12, p. 74–75]).

Remark 11. In the classical continuous-time case, all points are right-dense and consequently, Theorem 1 generalizes the usual continuous-time PMP where the maximization condition (9) is valid μ_L -almost everywhere (see [44, Theorem 6 p. 67]).

In the discrete-time setting, the possible failure of the maximization condition is a well known fact (see e.g. [12, p. 50–63]), and a fortiori in the time scale setting the maximization condition cannot be expected to hold in general at right-scattered points (see counterexamples below).

Many works have been devoted to derive a PMP in the discrete-time setting (see e.g. [7, 12, 22, 31, 37, 38, 43]). Since the maximization condition cannot be expected to hold true in general for discrete-time optimal control problems, it must be replaced with a weaker condition, of the

kind (8), involving the derivative of H with respect to u . Such a kind of inequality is provided in [12, Theorem 42.1 p. 330] for finite horizon problems and in [7] for infinite horizon problems. Our condition (8) is of a more general nature, as discussed next. In [37, 38, 46] the authors assume *directional convexity*, that is, for all $(v, v') \in \Omega^2$ and every $\theta \in [0, 1]$, there exists $v_\theta \in \Omega$ such that

$$f(q, v_\theta, t) = \theta f(q, v, t) + (1 - \theta)f(q, v', t), \quad f^0(q, v_\theta, t) \leq \theta f^0(q, v, t) + (1 - \theta)f^0(q, v', t),$$

for every $q \in \mathbb{R}^n$ and every $t \in \mathbb{T}$; and under this assumption they derive the maximization condition in the discrete-time case (see also [22] and [46, p. 235]). Note that this assumption is satisfied whenever Ω is convex, the dynamics f is affine with respect to u , and f^0 is convex in u (which implies that H is concave in u). We refer also to [43] where it is shown that, in the absence of such convexity assumptions, an approximate maximization condition can however be derived.

Note that, under additional assumptions, (8) implies the maximization condition. More precisely, let $r \in [a, b]_{\mathbb{T}} \cap \text{RS}$ and let $(q^*(\cdot), u^*(\cdot), p(\cdot), p^0)$ be the optimal extremal of Theorem 1. Let $r \in [a, b]_{\mathbb{T}} \cap \text{RS}$. If the function $u \mapsto H(q^*(r), u, p^\sigma(r), p^0, r)$ is concave on \mathbb{R}^m , then the inequality (8) implies that

$$H(q^*(r), u^*(r), p^\sigma(r), p^0, r) = \max_{v \in \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(u^*(r)))} H(q^*(r), v, p^\sigma(r), p^0, r).$$

If moreover $\Omega \subset \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(u^*(r)))$ (this is the case if Ω is convex), since $u^*(r) \in \Omega$, it follows that

$$H(q^*(r), u^*(r), p^\sigma(r), p^0, r) = \max_{v \in \Omega} H(q^*(r), v, p^\sigma(r), p^0, r).$$

Therefore, in particular, if H is concave in u and Ω is convex then the maximization condition holds as well at every right-scattered point.

Remark 12. It is interesting to note that, if H is convex in u then a certain minimization condition can be derived at every right-scattered point, as follows.

For every $v \in \Omega$, let $\text{Opp}(v) = \{2v - v' \mid v' \in \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v))\}$ denote the symmetric of $\overline{\text{Co}}(\mathcal{D}_{\text{stab}}^\Omega(v))$ with respect to the center v . It obviously follows from (8) that

$$\left\langle \frac{\partial H}{\partial u}(q^*(r), u^*(r), p^\sigma(r), p^0, r), v - u^*(r) \right\rangle_{\mathbb{R}^m} \geq 0, \quad (15)$$

for every $r \in [a, b]_{\mathbb{T}} \cap \text{RS}$ and every $v \in \text{Opp}(u^*(r))$. If H is convex in u on \mathbb{R}^m , then the inequality (15) implies that

$$H(q^*(r), u^*(r), p^\sigma(r), p^0, r) = \min_{v \in \text{Opp}(u^*(r))} H(q^*(r), v, p^\sigma(r), p^0, r) \quad (16)$$

We next provide several very simple examples illustrating the previous remarks.

Example 6. Here we give a counterexample showing that, although the final time is not fixed, the maximized Hamiltonian may not vanish.

Set $\mathbb{T} = \mathbb{N}$, $n = m = 1$, $f(q, u, t) = u$, $f^0(q, u, t) = 1$, $\Omega = [0, 1]$, $j = 2$, $g(q_1, q_2) = (q_1, q_2)$ and $S = \{0\} \times \{3/2\}$. The corresponding optimal control problem is the problem of steering the discrete-time control one-dimensional system $q(n+1) = q(n) + u(n)$ from $q(0) = 0$ to $q(b) = 3/2$ in minimal time, with control constraints $0 \leq u(k) \leq 1$. It is clear that the minimal time is $b^* = 2$, and that any control u such that $0 \leq u(0) \leq 1$, $0 \leq u(1) \leq 1$, and $u(0) + u(1) = 3/2$, is optimal.

Among these optimal controls, consider u^* defined by $u^*(0) = 1/2$ and $u^*(1) = 1$. Consider ψ , $p^0 \leq 0$ and $p(\cdot)$ the adjoint vector whose existence is asserted by the PMP. Since $u^*(0) \in \text{Int}(\Omega)$, it follows from (8) that $p(1) = 0$. The Hamiltonian is $H(q, u, p, p^0, t) = pu + p^0$, and since it is independent of q , it follows that $p(\cdot)$ is constant and thus equal to 0. In particular, $p(0) = p(2) = 0$ and hence $\psi = 0$. From the nontriviality condition $(p^0, \psi) \neq (0, 0)$ we infer that $p^0 \neq 0$. Therefore the maximized Hamiltonian at the final time is here equal to p^0 and thus is not equal to 0.

Example 7. Here we give a counterexample (in the spirit of [12, Examples 10.1-10.4 p. 59–62]) showing the failure of the maximization condition at right-scattered points.

Set $\mathbb{T} = \{0, 1, 2\}$, $n = m = 1$, $f(q, u, t) = u - q$, $f^0(q, u, t) = 2q^2 - u^2$, $\Omega = [0, 1]$, $j = 1$, $g(q_1, q_2) = q_1$ and $S = \{0\}$. Any solution of the resulting control system is such that $q(0) = 0$, $q(1) = u(0)$, $q(2) = u(1)$, and its cost is equal to $u(0)^2 - u(1)^2$. It follows that the optimal control u^* is unique and is such that $u^*(0) = 0$ and $u^*(1) = 1$. The Hamiltonian is $H(q, u, p, p^0, t) = p(u - q) + p^0(2q^2 - u^2)$. Consider ψ , $p^0 \leq 0$ and $p(\cdot)$ the adjoint vector whose existence is asserted by the PMP. Since g does not depend on q_2 , it follows that $p(2) = 0$, and from the extremal equations we infer that $p(1) = 0$ and $p(0) = 0$. Therefore $\psi = 0$ and hence $p^0 \neq 0$ (nontriviality condition) and we can assume that $p^0 = -1$. It follows that the maximized Hamiltonian is equal to $-p^0 = 1$ at $r = 0, 1, 2$, whereas $H(q^*(0), u^*(0), p(1), p^0, 0) = 0$. In particular, the maximization condition (9) is not satisfied at $r = 0 \in \text{RS}$ (note that it is however satisfied at $r = 1$).

Note that, in accordance with the fact that H is convex in u and $\text{Opp}(u^*(0)) =]-\infty, 0]$ and $\text{Opp}(u^*(1)) = [1, +\infty[$, the minimization condition (16) is indeed satisfied (see Remark 12).

Example 8. Here we give a counterexample in which, although the Hamiltonian is autonomous (independent of t), the maximized Hamiltonian is not constant over \mathbb{T} .

Set $\mathbb{T} = \{0, 1, 2\}$, $n = m = 1$, $f(q, u, t) = u - q$, $f^0(q, u, t) = (u^2 - q^2)/2$, $j = 1$, $g(q_1, q_2) = q_1$, $S = \{1\}$, $\Omega = [0, 1]$ and $b = 2$. Any solution of the resulting control system is such that $q(0) = 1$, $q(1) = u(0)$, $q(2) = u(1)$, and its cost is equal to $(u(1)^2 - 1)/2$. It follows that any control u such that $u(1) = 0$ is optimal (the value of $u(0)$ is arbitrary). Consider the optimal control u^* defined by $u^*(0) = u^*(1) = 0$, and let $q^*(\cdot)$ be the corresponding trajectory. Then $q^*(0) = 1$ and $q^*(1) = q^*(2) = 0$. The Hamiltonian is $H(q, u, p, p^0, t) = p(u - q) + p^0(u^2 - q^2)/2$. Consider ψ , $p^0 \leq 0$ and $p(\cdot)$ the adjoint vector whose existence is asserted by the PMP. Since g does not depend on q_2 , it follows that $p(2) = 0$, and from the extremal equations we infer that $p(1) = 0$ and $p(0) = -p^0$. In particular, from the nontriviality condition one has $p^0 \neq 0$ and we can assume that $p^0 = -1$. Therefore $H(q^*(0), v, p(1), p^0, 0) = 1/2 - v^2$ and $H(q^*(1), v, p(2), p^0, 1) = -v^2/2$, and it easily follows that the maximization condition holds at $r = 0$ and $r = 1$. This is in accordance with the fact that H is concave in u and Ω is convex. Moreover, the maximized Hamiltonian is equal to $1/2$ at $r = 0$, and to 0 at $r = 1$ and $r = 2$.

3 Proof of the main result

3.1 Preliminary comments

There exist several proofs of the continuous-time PMP in the literature. Mainly they can be classified as variants of two different approaches: the first of which consists of using a fixed point argument, and the second consists of using Ekeland's Variational Principle.

More precisely, the classical (and historical) proof of [44] relies on the use of the so-called needle-like variations combined with a fixed point Brouwer argument (see also [32, 40]). There exist variants, relying on the use of a conic version of the Implicit Function Theorem (see [4] or [30, 47]), the proof of which being however based on a fixed point argument. The proof of [17] uses a separation theorem (Hahn-Banach arguments) for cones combined with the Brouwer fixed point theorem. We could cite many other variants, all of them relying, at some step, on a fixed point argument.

The proof of [23] is of a different nature and follows from the combination of needle-like variations with Ekeland's Variational Principle. It does not rely on a fixed point argument. By the way note that this proof leads as well to an approximate PMP (see [23]), and withstands generalizations to the infinite dimensional setting (see e.g. [41]).

Note that, in all cases, needle-like variations are used to generate the so-called Pontryagin cone, serving as a first-order convex approximation of the reachable set. The adjoint vector is then constructed by propagating backward in time a Lagrange multiplier which is normal to this cone. Roughly, needle-like variations are kinds of perturbations of the reference control in L^1 topology (perturbations with arbitrary values, over small intervals of time) which generate perturbations of the trajectories in \mathcal{C}^0 topology.

Due to obvious topological obstructions, it is evident that the classical strategy of needle-like variations combined with a fixed point argument cannot hold in general in the time scale setting. At least one should distinguish between dense points and scattered points of \mathbb{T} . But even this distinction is not sufficient. Indeed, when applying the Brouwer fixed point Theorem to the mapping built on needle-like variations (see [40, 44]), it appears to be crucial that the domain of this mapping be convex. Roughly speaking, this domain consists of the product of the intervals of the spikes (intervals of perturbation). This requirement obviously excludes the scattered points of a time scale (which have anyway to be treated in another way), but even at some right-dense point $s \in \text{RD}$, there does not necessarily exist $\varepsilon > 0$ such that $[s, s + \varepsilon] \subset \mathbb{T}$. At such a point we can only ensure that 0 is not isolated in the set $\{\beta \geq 0 \mid s + \beta \in \mathbb{T}\}$. In our opinion this basic obstruction makes impossible the use of a fixed point argument in order to derive the PMP on a general time scale. Of course to overcome this difficulty one can assume that the μ_Δ -measure of right-dense points not admitting a right interval included in \mathbb{T} is zero. This assumption is however not very natural and would rule out time scales such as a generalized Cantor set having a positive μ_L -measure. Another serious difficulty that we are faced with on a general time scale is the technical fact that the formula (3), accounting for Lebesgue points, is valid only for β such that $s + \beta \in \mathbb{T}$. Actually if $s + \beta \notin \mathbb{T}$ then (3) is not true any more in general (it is very easy to construct a time scale \mathbb{T} for which (3) fails whenever $s + \beta \notin \mathbb{T}$, even with $q = 1$). Note that the concept of Lebesgue point is instrumental in the classical proof of the PMP in order to ensure that the needle-like variations can be built at different times¹ (see [40, 44]). On a general time scale this technical point would raise a serious issue².

The proof of the PMP that we provide in this article is based on Ekeland's Variational Principle, which permits to avoid the above obstructions and happens to be well adapted for the proof of a general PMP on time scales. It requires however the treatment of other kinds of technicalities, one of them being the concept of stable Ω -dense direction that we were led to introduce. Another point is that Ekeland's Variational Principle requires a complete metric space, which has led us to assume that Ω is closed (see Footnote 4).

Remark 13. Recall that a weak PMP (see Remark 10) on time scales is proved in [36] for shifted optimal control problems (see also [35]). A similar result can be derived in an analogous way for the non shifted optimal control problems (4) considered here. Since then, deriving the (strong) PMP on time scales was an open problem. While we were working on the contents of the present article (together with the companion paper [16]), at some step we discovered the publication of the article [49], in which the authors claim to have obtained a general version of the PMP. As in our work, their approach is based on Ekeland's Variational Principle. However, as already mentioned in the introduction, many arguments thereof are erroneous, and we believe that their errors cannot be corrected easily.

Although it is not our aim to be involved in controversy, since we were incidentally working in parallel on the same subject (deriving the PMP on time scales), we provide hereafter some evidence of the serious mistakes contained in [49].

¹More precisely, what is used in the approximate continuity property (see e.g. [24]).

²We are actually able to overcome this difficulty by considering multiple variations at right-scattered points, however this requires to assume that the set Ω is locally convex. The proof that we present further does not require such an assumption.

Note that, in order to derive a maximization condition Δ -almost everywhere (even at right-scattered points), as in [37, 38, 46] the authors of [49] assume directional convexity of the dynamics (see Remark 11 for the definition).

A first serious mistake in [49] is the fact that, in the application of Ekeland's Variational Principle, the authors use two different distances, depending on the nature of the point of \mathbb{T} under consideration (right-scattered or dense). As is usual in the proof of the PMP by Ekeland's Principle, the authors deduce from considerations on sequences of perturbation controls u^ε the existence of Lagrange multipliers φ_0 and ψ_0 ; the problem is that these multipliers are built separately for right-scattered and dense points (see [49, (32), (43), (52), (60)]), and thus are different in general since the distances used are different. Since the differential equation of the adjoint vector ψ depends on these multipliers, the existence of the adjoint vector in the main result [49, Theorem 3.1] cannot be established.

A second serious mistake is in the use of the directional convexity assumption (see [49, Equations (35), (36), (43)]). The first equality in (35) can obviously fail: the term $u_{\omega, \lambda}^\varepsilon(\tau)$ is a convex combination of $u^\varepsilon(\tau)$ and $u_{\mu(\tau)}^\varepsilon$ since $\mathcal{V}(\tau)$ is assumed to be convex, but the parameter of this convex combination is not necessarily equal to λ as claimed by the authors (unless f is affine in u and f^0 is convex in u , but this restrictive assumption is not made). The nasty consequence of this error is that, in (43), the limit as λ tends to 0 is not valid.

A third mistake is in [49, (57), (60), (62)], when the authors claim that the rest of the proof can be led for dense points similarly as for right-scattered points. They pass to the limit in (60) as ε tends to 0 and get that $V^\varepsilon(b)$ tends to $V(b)$, where V^ε is defined by (57) and V is defined similarly. However, this does not hold true. Indeed, even though $d_\Delta^*(u^\varepsilon, u^*)$ (Ekeland's distance) tends to 0, there is no guarantee that $u^\varepsilon(\tau)$ tends to $u^*(\tau)$.

The above mistakes are major and cannot be corrected even through a major revision of the overall proof, due to evident obstructions. There are many other minor ones along the paper (which can be corrected, although some of them require a substantial work), such as: the Δ -measurability of the map \mathcal{V} is not proved; in (45) the authors should consider subsequences and not a global limit; in (55), any arbitrary $\rho > 0$ cannot be considered to deal with the Δ -Lebesgue point τ , but only with $\tau - \rho \in \mathbb{T}$ (recall that the equality (3) of our paper is valid only if $s + \beta \in \mathbb{T}$, and that, as already mentioned, on a general time scale Lebesgue points must be handled with some special care).

In view of these numerous issues, it cannot be considered that the PMP has been proved in [49]. The aim of the present article (whose work was initiated far before we discovered the publication [49]) is to fill a gap in the literature and to derive a general strong version of the PMP on time scales. Finally, it can be noted that the authors of [49] make restrictive assumptions: their set Ω is convex and is compact at scattered points, their dynamics are globally Lipschitzian and directionally convex, and they consider optimal control problems with fixed final time and fixed initial and final points. In the present article we go far beyond these unnecessary and not natural requirements, as already explained throughout.

3.2 Needle-like variations of admissible controls

Let $b \in \mathbb{T} \setminus \{a\}$. Following the definition of an admissible control (see Definition 2 in Section 2.3), we denote by $\mathcal{U}\mathcal{Q}_{\text{ad}}^b$ the set of all $(u, q_a) \in \mathcal{U} \times \mathbb{R}^n$ such that u is an admissible control on $[a, b]_{\mathbb{T}}$ associated with the initial condition q_a . It is endowed with the distance

$$d_{\mathcal{U}\mathcal{Q}_{\text{ad}}^b}((u, q_a), (u', q'_a)) = \|u - u'\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}, \mathbb{R}^m)} + \|q_a - q'_a\|_{\mathbb{R}^n}. \quad (17)$$

Throughout the section, we consider $(u, q_a) \in \mathcal{U}\mathcal{Q}_{\text{ad}}^b$ with $u \in L_{\mathbb{T}}^\infty([a, b]_{\mathbb{T}}, \Omega)$ and the corresponding solution $q(\cdot, u, q_a)$ of (4) with $q(a) = q_a$. This section 3.2 is devoted to define appropriate

variations of (u, q_a) , instrumental in order to prove the PMP. We present some preliminary topological results in Section 3.2.1. Then we define needle-like variations of u in Sections 3.2.2 and 3.2.3, respectively at a right-scattered point and at a right-dense point and derive some useful properties. Finally in Section 3.2.4 we make some variations of the initial condition q_a .

3.2.1 Preliminaries

In the first lemma below, we prove that $\mathcal{UQ}_{\text{ad}}^b$ is open. Actually we prove a stronger result, by showing that $\mathcal{UQ}_{\text{ad}}^b$ contains a neighborhood of any of its point in L^1 topology, which will be useful in order to define needle-like variations.

Lemma 1. *Let $R > \|u\|_{L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}, \mathbb{R}^m)}$. There exist $\nu_R > 0$ and $\eta_R > 0$ such that the set*

$$\begin{aligned} E(u, q_a, R) = \{ (u', q'_a) \in \mathcal{U} \times \mathbb{R}^n \mid & \|u' - u\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}, \mathbb{R}^m)} \leq \nu_R, \\ & \|u'\|_{L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}, \mathbb{R}^m)} \leq R, \|q'_a - q_a\|_{\mathbb{R}^n} \leq \eta_R \} \end{aligned}$$

is contained in $\mathcal{UQ}_{\text{ad}}^b$.

Before proving this lemma, let us recall a time scale version of Gronwall's Lemma (see [10, Chapter 6.1]). The generalized exponential function is defined by $e_L(t, c) = \exp(\int_{[c, t]_{\mathbb{T}}} \xi_{\mu(\tau)}(L) \Delta\tau)$, for every $L \geq 0$, every $c \in \mathbb{T}$ and every $t \in [c, +\infty]_{\mathbb{T}}$, where $\xi_{\mu(\tau)}(L) = \log(1 + L\mu(\tau))/\mu(\tau)$ whenever $\mu(\tau) > 0$, and $\xi_{\mu(\tau)}(L) = L$ whenever $\mu(\tau) = 0$ (see [10, Chapter 2.2]). Note that, for every $L \geq 0$ and every $c \in \mathbb{T}$, the function $e_L(\cdot, c)$ (resp. $e_L(c, \cdot)$) is positive and increasing on $[c, +\infty]_{\mathbb{T}}$ (resp. positive and decreasing on $[a, c]_{\mathbb{T}}$), and moreover there holds $e_L(t_2, t_1)e_L(t_1, c) = e_L(t_2, c)$, for every $L \geq 0$ and all $(c, t_1, t_2) \in \mathbb{T}^3$ such that $c \leq t_1 \leq t_2$.

Lemma 2 ([10]). *Let $(c, d) \in \mathbb{T}^2$ such that $c < d$, let L_1 and L_2 be nonnegative real numbers, and let $q \in \mathcal{C}([c, d]_{\mathbb{T}}, \mathbb{R})$ satisfying $0 \leq q(t) \leq L_1 + L_2 \int_{[c, t]_{\mathbb{T}}} q(\tau) \Delta\tau$, for every $t \in [c, d]_{\mathbb{T}}$. Then $0 \leq q(t) \leq L_1 e_{L_2}(t, c)$, for every $t \in [c, d]_{\mathbb{T}}$.*

Proof of Lemma 1. Let $R > \|u\|_{L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}, \mathbb{R}^m)}$. By continuity of $q(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$, the set

$$K = \{(x, v, t) \in \mathbb{R}^n \times \overline{B}_{\mathbb{R}^m}(0, R) \times [a, b]_{\mathbb{T}} \mid \|x - q(t, u, q_a)\|_{\mathbb{R}^n} \leq 1\}$$

is a compact subset of $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{T}$. Therefore $\|\partial f / \partial q\|$ and $\|\partial f / \partial u\|$ are bounded by some $L \geq 0$ on K and moreover L is chosen such that

$$\|f(x_1, v_1, t) - f(x_2, v_2, t)\|_{\mathbb{R}^n} \leq L(\|x_1 - x_2\|_{\mathbb{R}^n} + \|v_1 - v_2\|_{\mathbb{R}^m}), \quad (18)$$

for all (x_1, v_1, t) and (x_2, v_2, t) in K . Let $\nu_R > 0$ and $0 < \eta_R < 1$ such that $(\eta_R + \nu_R L)e_L(b, a) < 1$. Note that K , L , ν_R and η_R depend on (u, q_a, R) .

Let $(u', q'_a) \in E(u, q_a, R)$. We denote by $I_{\mathbb{T}}'$ the interval of definition of $q(\cdot, u', q'_a)$ satisfying $a \in I_{\mathbb{T}}'$ and $I_{\mathbb{T}}' \setminus \{a\} \neq \emptyset$. It suffices to prove that $b \in I_{\mathbb{T}}'$. By contradiction, assume that the set $A = \{t \in I_{\mathbb{T}}' \cap [a, b]_{\mathbb{T}} \mid \|q(t, u', q'_a) - q(t, u, q_a)\|_{\mathbb{R}^n} > 1\}$ is not empty and set $t_1 = \inf A$. Since \mathbb{T} is closed, $t_1 \in I_{\mathbb{T}}' \cap [a, b]_{\mathbb{T}}$ and $[a, t_1]_{\mathbb{T}} \subset I_{\mathbb{T}}' \cap [a, b]_{\mathbb{T}}$. If t_1 is a minimum then $\|q(t_1, u', q'_a) - q(t_1, u, q_a)\|_{\mathbb{R}^n} > 1$. If t_1 is not a minimum then $t_1 \in \text{RD}$ and by continuity we have $\|q(t_1, u', q'_a) - q(t_1, u, q_a)\|_{\mathbb{R}^n} \geq 1$. Moreover there holds $t_1 > a$ since $\|q(a, u', q'_a) - q(a, u, q_a)\|_{\mathbb{R}^n} = \|q'_a - q_a\|_{\mathbb{R}^n} \leq \eta_R < 1$. Hence $\|q(\tau, u', q'_a) - q(\tau, u, q_a)\|_{\mathbb{R}^n} \leq 1$ for every $\tau \in [a, t_1]_{\mathbb{T}}$. Therefore $(q(\tau, u', q'_a), u'(\tau), \tau)$ and $(q(\tau, u, q_a), u(\tau), \tau)$ are elements of K for Δ -a.e. $\tau \in [a, t_1]_{\mathbb{T}}$. Since there holds

$$q(t, u', q'_a) - q(t, u, q_a) = q'_a - q_a + \int_{[a, t]_{\mathbb{T}}} (f(q(\tau, u', q'_a), u'(\tau), \tau) - f(q(\tau, u, q_a), u(\tau), \tau)) \Delta\tau,$$

for every $t \in I'_\mathbb{T} \cap [a, b]_\mathbb{T}$, it follows from (18) and from Lemma 2 that, for every $t \in [a, t_1]_\mathbb{T}$,

$$\begin{aligned} \|q(t, u', q'_a) - q(t, u, q_a)\|_{\mathbb{R}^n} &\leq \|q'_a - q_a\|_{\mathbb{R}^n} + L \int_{[a, t]_\mathbb{T}} \|u'(\tau) - u(\tau)\|_{\mathbb{R}^m} \Delta\tau \\ &\quad + L \int_{[a, t]_\mathbb{T}} \|q(\tau, u', q'_a) - q(\tau, u, q_a)\|_{\mathbb{R}^n} \Delta\tau \\ &\leq (\|q'_a - q_a\|_{\mathbb{R}^n} + L\|u' - u\|_{L^1_\mathbb{T}([a, b]_\mathbb{T}, \mathbb{R}^m)})e_L(b, a) \\ &\leq (\eta_R + \nu_R L)e_L(b, a) < 1. \end{aligned}$$

This raises a contradiction at $t = t_1$. Therefore A is empty and thus $q(\cdot, u', q'_a)$ is bounded on $I'_\mathbb{T} \cap [a, b]_\mathbb{T}$. It follows from [16, Theorem 2] that $b \in I'_\mathbb{T}$, that is, $(u', q'_a) \in \mathcal{UQ}_{\text{ad}}^b$. \square

Remark 14. Let $(u', q'_a) \in E(u, q_a, R)$. With the notations of the above proof, since $I'_\mathbb{T} \cap [a, b]_\mathbb{T} = [a, b]_\mathbb{T}$ and A is empty, we infer that $\|q(t, u', q'_a) - q(t, u, q_a)\| \leq 1$, for every $t \in [a, b]_\mathbb{T}$. Therefore $(q(t, u', q'_a), u'(t), t) \in K$ for every $(u', q'_a) \in E(u, q_a, R)$ and for Δ -a.e. $t \in [a, b]_\mathbb{T}$.

Lemma 3. *With the notations of Lemma 1, the mapping*

$$\begin{aligned} F_{(u, q_a, R)} : (E(u, q_a, R), d_{\mathcal{UQ}_{\text{ad}}^b}) &\longrightarrow (\mathcal{C}([a, b]_\mathbb{T}, \mathbb{R}^n), \|\cdot\|_\infty) \\ (u', q'_a) &\longmapsto q(\cdot, u', q'_a) \end{aligned}$$

is Lipschitzian. In particular, for every $(u', q'_a) \in E(u, q_a, R)$, $q(\cdot, u', q'_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_\mathbb{T}$ when u' tends to u in $L^1_\mathbb{T}([a, b]_\mathbb{T}, \mathbb{R}^m)$ and q'_a tends to q_a in \mathbb{R}^n .

Proof. Let (u', q'_a) and (u'', q''_a) be elements of $E(u, q_a, R) \subset \mathcal{UQ}_{\text{ad}}^b$. It follows from Remark 14 that $(q(\tau, u'', q''_a), u''(\tau), \tau)$ and $(q(\tau, u', q'_a), u'(\tau), \tau)$ are elements of K for Δ -a.e. $\tau \in [a, b]_\mathbb{T}$. Following the same arguments as in the previous proof, it follows from (18) and from Lemma 2 that, for every $t \in [a, b]_\mathbb{T}$,

$$\|q(t, u'', q''_a) - q(t, u', q'_a)\|_{\mathbb{R}^n} \leq (\|q''_a - q'_a\|_{\mathbb{R}^n} + L\|u'' - u'\|_{L^1_\mathbb{T}([a, b]_\mathbb{T}, \mathbb{R}^m)})e_L(b, a).$$

The lemma follows. \square

3.2.2 Needle-like variation of u at a right-scattered point

Let $r \in [a, b]_\mathbb{T} \cap \text{RS}$ and let $y \in \mathcal{D}^\Omega(u(r))$. We define the needle-like variation $\Pi = (r, y)$ of u at the right-scattered point r by

$$u_\Pi(t, \alpha) = \begin{cases} u(r) + \alpha(y - u(r)) & \text{if } t = r, \\ u(t) & \text{if } t \neq r. \end{cases}$$

for every $\alpha \in \mathcal{D}^\Omega(u(r), y)$. It follows from Section 2.2 that $u_\Pi(\cdot, \alpha) \in L^\infty_\mathbb{T}([a, b]_\mathbb{T}; \Omega)$.

Lemma 4. *There exists $\alpha_0 > 0$ such that $(u_\Pi(\cdot, \alpha), q_a) \in \mathcal{UQ}_{\text{ad}}^b$, for every $\alpha \in \mathcal{D}^\Omega(u(r), y) \cap [0, \alpha_0]$.*

Proof. Let $R = \max(\|u\|_{L^\infty_\mathbb{T}([a, b]_\mathbb{T}, \mathbb{R}^m)}, \|u(r)\|_{\mathbb{R}^m} + \|y\|_{\mathbb{R}^m}) + 1 > \|u\|_{L^\infty_\mathbb{T}([a, b]_\mathbb{T}, \mathbb{R}^m)}$. We use the notations K , L , ν_R and η_R , associated with (u, q_a, R) , defined in Lemma 1 and in its proof.

One has $\|u_\Pi(\cdot, \alpha)\|_{L^\infty_\mathbb{T}([a, b]_\mathbb{T}, \mathbb{R}^m)} \leq R$ for every $\alpha \in \mathcal{D}^\Omega(u(r), y)$, and

$$\|u_\Pi(\cdot, \alpha) - u\|_{L^1_\mathbb{T}([a, b]_\mathbb{T}, \mathbb{R}^m)} = \mu(r)\|u_\Pi(r, \alpha) - u(r)\|_{\mathbb{R}^m} = \alpha\mu(r)\|y - u(r)\|_{\mathbb{R}^m}.$$

Hence, there exists $\alpha_0 > 0$ such that $\|u_\Pi(\cdot, \alpha) - u\|_{L^1_\mathbb{T}([a, b]_\mathbb{T}, \mathbb{R}^m)} \leq \nu_R$ for every $\alpha \in \mathcal{D}^\Omega(u(r), y) \cap [0, \alpha_0]$, and hence $(u_\Pi(\cdot, \alpha), q_a) \in E(u, q_a, R)$. The claim follows then from Lemma 1. \square

Lemma 5. *The mapping*

$$\begin{aligned} F_{(u, q_a, \Pi)} : (\mathcal{D}^\Omega(u(r), y) \cap [0, \alpha_0], |\cdot|) &\longrightarrow (\mathcal{C}([a, b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_\infty) \\ \alpha &\longmapsto q(\cdot, u_\Pi(\cdot, \alpha), q_a) \end{aligned}$$

is Lipschitzian. In particular, for every $\alpha \in \mathcal{D}^\Omega(u(r), y) \cap [0, \alpha_0]$, $q(\cdot, u_\Pi(\cdot, \alpha), q_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$ as α tends to 0.

Proof. We use the notations of proof of Lemma 4. It follows from Lemma 3 that there exists $C \geq 0$ (the Lipschitz constant of $F_{(u, q_a, R)}$) such that

$$\begin{aligned} \|q(\cdot, u_\Pi(\cdot, \alpha^2), q_a) - q(\cdot, u_\Pi(\cdot, \alpha^1), q_a)\|_\infty &\leq C d_{\mathcal{U}_{\mathcal{Q}_{\text{ad}}^b}}((u_\Pi(\cdot, \alpha^2), q_a), (u_\Pi(\cdot, \alpha^1), q_a)) \\ &= C |\alpha^2 - \alpha^1| \mu(r) \|y - u(r)\|_{\mathbb{R}^m}, \end{aligned}$$

for all α^1 and α^2 in $\mathcal{D}^\Omega(u(r), y) \cap [0, \alpha_0]$. The lemma follows. \square

We define the so-called *variation vector* $w_\Pi(\cdot, u, q_a)$ associated with the needle-like variation $\Pi = (r, y)$ as the unique solution on $[\sigma(r), b]_{\mathbb{T}}$ of the linear Δ -Cauchy problem

$$w^\Delta(t) = \frac{\partial f}{\partial q}(q(t, u(t), q_a), u(t), t) w(t), \quad w(\sigma(r)) = \mu(r) \frac{\partial f}{\partial u}(q(r, u, q_a), u(r), r)(y - u(r)). \quad (19)$$

The existence and uniqueness of $w_\Pi(\cdot, u, q_a)$ are ensured by [16, Theorem 3].

Proposition 1. *The mapping*

$$\begin{aligned} F_{(u, q_a, \Pi)} : (\mathcal{D}^\Omega(u(r), y) \cap [0, \alpha_0], |\cdot|) &\longrightarrow (\mathcal{C}([\sigma(r), b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_\infty) \\ \alpha &\longmapsto q(\cdot, u_\Pi(\cdot, \alpha), q_a) \end{aligned} \quad (20)$$

is differentiable³ at 0, and there holds $DF_{(u, q_a, \Pi)}(0) = w_\Pi(\cdot, u, q_a)$.

Proof. We use the notations of proof of Lemma 4. Recall that $(q(t, u_\Pi(\cdot, \alpha), q_a), u_\Pi(t, \alpha), t) \in K$ for every $\alpha \in \mathcal{D}^\Omega(u(r), y) \cap [0, \alpha_0]$ and for Δ -a.e. $t \in [a, b]_{\mathbb{T}}$, see Remark 14. For every $\alpha \in \mathcal{D}^\Omega(u(r), y) \cap [0, \alpha_0]$ and every $t \in [\sigma(r), b]_{\mathbb{T}}$, we define

$$\varepsilon_\Pi(t, \alpha) = \frac{q(t, u_\Pi(\cdot, \alpha), q_a) - q(t, u, q_a)}{\alpha} - w_\Pi(t, u, q_a).$$

It suffices to prove that $\varepsilon_\Pi(\cdot, \alpha)$ converges uniformly to 0 on $[\sigma(r), b]_{\mathbb{T}}$ as α tends to 0. For every $\alpha \in \mathcal{D}^\Omega(u(r), y) \cap [0, \alpha_0]$, the function $\varepsilon_\Pi(\cdot, \alpha)$ is absolutely continuous on $[\sigma(r), b]_{\mathbb{T}}$, and $\varepsilon_\Pi(t, \alpha) = \varepsilon_\Pi(\sigma(r), \alpha) + \int_{[\sigma(r), t]_{\mathbb{T}}} \varepsilon_\Pi^\Delta(\tau, \alpha) \Delta\tau$ for every $t \in [\sigma(r), b]_{\mathbb{T}}$, where

$$\varepsilon_\Pi^\Delta(t, \alpha) = \frac{f(q(t, u_\Pi(\cdot, \alpha), q_a), u(t), t) - f(q(t, u, q_a), u(t), t))}{\alpha} - \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t) w_\Pi(t, u, q_a),$$

for Δ -a.e. $t \in [\sigma(r), b]_{\mathbb{T}}$. It follows from the Mean Value Theorem applied for Δ -a.e. $t \in [\sigma(r), b]_{\mathbb{T}}$ to the function defined by $\varphi_t(\theta) = f((1 - \theta)q(t, u, q_a) + \theta q(t, u_\Pi(\cdot, \alpha), q_a), u(t), t)$ for every $\theta \in [0, 1]$, that there exists $\theta_\Pi(t, \alpha) \in \mathbb{R}^n$, belonging to the segment of extremities $q(t, u, q_a)$ and $q(t, u_\Pi(\cdot, \alpha), q_a)$, such that

$$\begin{aligned} \varepsilon_\Pi^\Delta(t, \alpha) &= \frac{\partial f}{\partial q}(\theta_\Pi(t, \alpha), u(t), t) \varepsilon_\Pi(t, \alpha) \\ &\quad + \left(\frac{\partial f}{\partial q}(\theta_\Pi(t, \alpha), u(t), t) - \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t) \right) w_\Pi(t, u, q_a). \end{aligned}$$

³Clearly this mapping can be extended to a neighborhood of 0 and we speak of its differential at 0 in this sense.

Since $(\theta_\Pi(t, \alpha), u(t), t) \in K$ for Δ -a.e. $t \in [\sigma(r), b]_\mathbb{T}$, it follows that $\|\varepsilon_\Pi^\Delta(t, \alpha)\| \leq \chi_\Pi(t, \alpha) + L\|\varepsilon_\Pi(t, \alpha)\|$, where $\chi_\Pi(t, \alpha) = \left\| \left(\frac{\partial f}{\partial q}(\theta_\Pi(t, \alpha), u(t), t) - \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t) \right) w_\Pi(t, u, q_a) \right\|$. Therefore, one has

$$\|\varepsilon_\Pi(t, \alpha)\|_{\mathbb{R}^n} \leq \|\varepsilon_\Pi(\sigma(r), \alpha)\|_{\mathbb{R}^n} + \int_{[\sigma(r), b]_\mathbb{T}} \chi_\Pi(\tau, \alpha) \Delta\tau + L \int_{[\sigma(r), t]_\mathbb{T}} \|\varepsilon_\Pi(\tau, \alpha)\|_{\mathbb{R}^n} \Delta\tau,$$

for every $t \in [\sigma(r), b]_\mathbb{T}$. It follows from Lemma 2 that $\|\varepsilon_\Pi(t, \alpha)\|_{\mathbb{R}^n} \leq \Upsilon_\Pi(\alpha)e_L(b, \sigma(r))$, for every $t \in [\sigma(r), b]_\mathbb{T}$, where $\Upsilon_\Pi(\alpha) = \|\varepsilon_\Pi(\sigma(r), \alpha)\|_{\mathbb{R}^n} + \int_{[\sigma(r), b]_\mathbb{T}} \chi_\Pi(\tau, \alpha) \Delta\tau$.

To conclude, it remains to prove that $\Upsilon_\Pi(\alpha)$ converges to 0 as α tends to 0. First, since $\theta_\Pi(\cdot, \alpha)$ converges uniformly to $q(\cdot, u, q_a)$ on $[\sigma(r), b]_\mathbb{T}$ as α tends to 0, and since $\partial f/\partial q$ is uniformly continuous on K , we infer that $\int_{[\sigma(r), b]_\mathbb{T}} \chi_\Pi(\tau, \alpha) \Delta\tau$ converges to 0 as α tends to 0. Second, it is easy to see that $\|\varepsilon_\Pi(\sigma(r), \alpha)\|_{\mathbb{R}^n}$ converges to 0 as α tends to 0. The conclusion follows. \square

Lemma 6. *Let $R > \|u\|_{L^\infty([a, b]_\mathbb{T}, \mathbb{R}^m)}$ and let $(u_k, q_{a,k})_{k \in \mathbb{N}}$ be a sequence of elements of $E(u, q_a, R)$. If u_k converges to u Δ -a.e. on $[a, b]_\mathbb{T}$ and $q_{a,k}$ converges to q_a in \mathbb{R}^n as k tends to $+\infty$, then $w_\Pi(\cdot, u_k, q_{a,k})$ converges uniformly to $w_\Pi(\cdot, u, q_a)$ on $[\sigma(r), b]_\mathbb{T}$ as k tends to $+\infty$.*

Proof. We use the notations K , L , ν_R and η_R , associated with (u, q_a, R) , defined in Lemma 1 and in its proof.

Consider the absolutely continuous function defined by $\Phi_k(t) = w_\Pi(t, u_k, q_{a,k}) - w_\Pi(t, u, q_a)$ for every $k \in \mathbb{N}$ and every $t \in [\sigma(r), b]_\mathbb{T}$. Let us prove that Φ_k converges uniformly to 0 on $[\sigma(r), b]_\mathbb{T}$ as k tends to $+\infty$. One has

$$\begin{aligned} \Phi_k(t) &= \Phi_k(\sigma(r)) + \int_{[\sigma(r), t]_\mathbb{T}} \frac{\partial f}{\partial q}(q(\tau, u_k, q_{a,k}), u_k(\tau), \tau) \Phi_k(\tau) \Delta\tau \\ &\quad + \int_{[\sigma(r), t]_\mathbb{T}} \left(\frac{\partial f}{\partial q}(q(\tau, u_k, q_{a,k}), u_k(\tau), \tau) - \frac{\partial f}{\partial q}(q(\tau, u, q_a), u(\tau), \tau) \right) w_\Pi(\tau, u, q_a) \Delta\tau, \end{aligned}$$

for every $t \in [\sigma(r), b]_\mathbb{T}$ and every $k \in \mathbb{N}$. Since $(u_k, q_{a,k}) \in E(u, q_a, R)$ for every $k \in \mathbb{N}$, it follows from Remark 14 that $(q(t, u_k, q_{a,k}), u_k(t), t) \in K$ and $(q(t, u, q_a), u(t), t) \in K$ for Δ -a.e. $t \in [a, b]_\mathbb{T}$. Hence it follows from Lemma 2 that

$$\|\Phi_k(t)\|_{\mathbb{R}^n} \leq (\|\Phi_k(\sigma(r))\|_{\mathbb{R}^n} + \vartheta_k)e_L(b, \sigma(r)),$$

for every $t \in [\sigma(r), b]_\mathbb{T}$, where

$$\vartheta_k = \int_{[\sigma(r), b]_\mathbb{T}} \left\| \frac{\partial f}{\partial q}(q(\tau, u_k, q_{a,k}), u_k(\tau), \tau) - \frac{\partial f}{\partial q}(q(\tau, u, q_a), u(\tau), \tau) \right\|_{\mathbb{R}^{n,n}} \|w_\Pi(\tau, u, q_a)\|_{\mathbb{R}^n} \Delta\tau.$$

Since $\mu_\Delta(\{r\}) = \mu(r) > 0$, $u_k(r)$ converges to $u(r)$ as k tends to $+\infty$. Moreover, $(u_k, q_{a,k})$ converges to (u, q_a) in $(E(u, q_a, R), d_{\mathcal{U}_{\mathcal{Q}_{\text{ad}}^b}})$ and, from Lemma 3, $q(\cdot, u_k, q_{a,k})$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_\mathbb{T}$ as k tends to $+\infty$. We infer that $\Phi_k(\sigma(r))$ converges to 0 as k tends to $+\infty$, and from the Lebesgue dominated convergence theorem we conclude that ϑ_k converges to 0 as k tends to $+\infty$. The lemma follows. \square

Remark 15. It is interesting to note that, since $u_k(r)$ converges to $u(r)$ as k tends to $+\infty$, if we assume that $y \in \mathcal{D}_{\text{stab}}^\Omega(u(r))$, then $y \in \mathcal{D}^\Omega(u_k(r))$ for k sufficiently large.

3.2.3 Needle-like variation of u at a right-dense point

The definition of a needle-like variation at a Lebesgue right-dense point is very similar to the classical continuous-time case. Let $s \in \mathcal{L}_{[a, b]_{\mathbb{T}}}(f(q(\cdot, u, q_a), u(\cdot, \cdot))) \cap \text{RD}$ and $z \in \Omega$. We define the needle-like variation $\Pi = (s, z)$ of u at s by

$$u_{\Pi}(t, \beta) = \begin{cases} z & \text{if } t \in [s, s + \beta]_{\mathbb{T}}, \\ u(t) & \text{if } t \notin [s, s + \beta]_{\mathbb{T}}. \end{cases}$$

for every $\beta \in \mathcal{V}_s^b$ (here, we use the notations introduced in Section 2.1). Note that $u_{\Pi}(\cdot, \beta) \in L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}; \Omega)$.

Lemma 7. *There exists $\beta_0 > 0$ such that $(u_{\Pi}(\cdot, \beta), q_a) \in \mathcal{U}_{\text{ad}}^b$ for every $\beta \in \mathcal{V}_s^b \cap [0, \beta_0]$.*

Proof. Let $R = \max(\|u\|_{L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}, \mathbb{R}^m)}, \|z\|_{\mathbb{R}^m}) + 1 > \|u\|_{L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}, \mathbb{R}^m)}$. We use the notations K, L, ν_R and η_R , associated with (u, q_a, R) , defined in Lemma 1 and in its proof.

For every $\beta \in \mathcal{V}_s^b$ one has $\|u_{\Pi}(\cdot, \beta)\|_{L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}, \mathbb{R}^m)} \leq R$ and

$$\|u_{\Pi}(\cdot, \beta) - u\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}, \mathbb{R}^m)} = \int_{[s, s+\beta]_{\mathbb{T}}} \|z - u(\tau)\|_{\mathbb{R}^m} \Delta\tau \leq 2R\beta.$$

Hence, there exists $\beta_0 > 0$ such that for every $\beta \in \mathcal{V}_s^b \cap [0, \beta_0]$, $\|u_{\Pi}(\cdot, \beta) - u\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}, \mathbb{R}^m)} \leq \nu_R$ and thus $(u_{\Pi}(\cdot, \beta), q_a) \in E(u, q_a, R)$. The conclusion then follows from Lemma 1. \square

Lemma 8. *The mapping*

$$\begin{aligned} F_{(u, q_a, \Pi)} : (\mathcal{V}_s^b \cap [0, \beta_0], |\cdot|) &\longrightarrow (\mathcal{C}([a, b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_{\infty}) \\ \beta &\longmapsto q(\cdot, u_{\Pi}(\cdot, \beta), q_a) \end{aligned}$$

is Lipschitzian. In particular, for every $\beta \in \mathcal{V}_s^b \cap \overline{B}(0, \beta_0)$, $q(\cdot, u_{\Pi}(\cdot, \beta), q_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$ as β tends to 0.

Proof. We use the notations of proof of Lemma 7. From Lemma 3, there exists $C \geq 0$ (Lipschitz constant of $F_{(u, q_a, R)}$) such that

$$\begin{aligned} \|q(\cdot, u_{\Pi}(\cdot, \beta^2), q_a) - q(\cdot, u_{\Pi}(\cdot, \beta^1), q_a)\|_{\infty} &\leq C d_{\mathcal{U}_{\text{ad}}^b}((u_{\Pi}(\cdot, \beta^2), q_a), (u_{\Pi}(\cdot, \beta^1), q_a)) \\ &\leq 2CR|\beta^2 - \beta^1|, \end{aligned}$$

for all β^1 and β^2 in $\mathcal{V}_s^b \cap [0, \beta_0]$. The lemma follows. \square

According to [16, Theorem 3], we define the *variation vector* $w_{\Pi}(\cdot, u, q_a)$ associated with the needle-like variation $\Pi = (s, z)$ as the unique solution on $[s, b]_{\mathbb{T}}$ of the linear Δ -Cauchy problem

$$w^{\Delta}(t) = \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t)w(t), \quad w(s) = f(q(s, u, q_a), z, s) - f(q(s, u, q_a), u(s), s). \quad (21)$$

Proposition 2. *For every $\delta \in \mathcal{V}_s^b \setminus \{0\}$, the mapping*

$$\begin{aligned} F_{(u, q_a, \Pi)}^{\delta} : (\mathcal{V}_s^b \cap [0, \beta_0], |\cdot|) &\longrightarrow (\mathcal{C}([s + \delta, b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_{\infty}) \\ \beta &\longmapsto q(\cdot, u_{\Pi}(\cdot, \beta), q_a) \end{aligned} \quad (22)$$

is differentiable at 0, and one has $DF_{(u, q_a, \Pi)}^{\delta}(0) = w_{\Pi}(\cdot, u, q_a)$.

Proof. We use the notations of proof of Lemma 7. Recall that $(q(t, u_\Pi(\cdot, \beta), q_a), u_\Pi(t, \beta), t)$ and $(q(t, u_\Pi(\cdot, \beta), q_a), z, t)$ belong to K for every $\beta \in \mathcal{V}_s^b \cap [0, \beta_0]$ and for Δ -a.e. $t \in [a, b]_\mathbb{T}$, see Remark 14. For every $\beta \in \mathcal{V}_s^b \cap [0, \beta_0]$ and every $t \in [s + \beta, b]_\mathbb{T}$, we define

$$\varepsilon_\Pi(t, \beta) = \frac{q(t, u_\Pi(\cdot, \beta), q_a) - q(t, u, q_a)}{\beta} - w_\Pi(t, u, q_a).$$

It suffices to prove that $\varepsilon_\Pi(\cdot, \beta)$ converges uniformly to 0 on $[s + \beta, b]_\mathbb{T}$ as β tends to 0 (note that, for every $\delta \in \mathcal{V}_s^b \setminus \{0\}$, it suffices to consider $\beta \leq \delta$). For every $\beta \in \mathcal{V}_s^b \cap [0, \beta_0]$, the function $\varepsilon_\Pi(\cdot, \beta)$ is absolutely continuous on $[s + \beta, b]_\mathbb{T}$ and $\varepsilon_\Pi(t, \beta) = \varepsilon_\Pi(s + \beta, \beta) + \int_{[s + \beta, t]_\mathbb{T}} \varepsilon_\Pi^\Delta(\tau, \beta) \Delta\tau$, for every $t \in [s + \beta, b]_\mathbb{T}$, where

$$\varepsilon_\Pi^\Delta(t, \beta) = \frac{f(q(t, u_\Pi(\cdot, \beta), q_a), u(t), t) - f(q(t, u, q_a), u(t), t))}{\beta} - \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t) w_\Pi(t, u, q_a).$$

for Δ -a.e. $t \in [s + \beta, b]_\mathbb{T}$. As in the proof of Proposition 1, it follows from the Mean Value Theorem that, for Δ -a.e. $t \in [s + \beta, b]_\mathbb{T}$, there exists $\theta_\Pi(t, \beta) \in \mathbb{R}^n$, belonging to the segment of extremities $q(t, u, q_a)$ and $q(t, u_\Pi(\cdot, \beta), q_a)$, such that

$$\begin{aligned} \varepsilon_\Pi^\Delta(t, \beta) &= \frac{\partial f}{\partial q}(\theta_\Pi(t, \beta), u(t), t) \varepsilon_\Pi(t, \beta) \\ &\quad + \left(\frac{\partial f}{\partial q}(\theta_\Pi(t, \beta), u(t), t) - \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t) \right) w_\Pi(t, u, q_a). \end{aligned}$$

Since $(\theta_\Pi(t, \beta), u(t), t) \in K$ for Δ -a.e. $t \in [s + \beta, b]_\mathbb{T}$, it follows that $\|\varepsilon_\Pi^\Delta(t, \beta)\| \leq \chi_\Pi(t, \beta) + L\|\varepsilon_\Pi(t, \beta)\|$, where $\chi_\Pi(t, \beta) = \left\| \left(\frac{\partial f}{\partial q}(\theta_\Pi(t, \beta), u(t), t) - \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t) \right) w_\Pi(t, u, q_a) \right\|$. Therefore, one has

$$\|\varepsilon_\Pi(t, \beta)\|_{\mathbb{R}^n} \leq \|\varepsilon_\Pi(s + \beta, \beta)\|_{\mathbb{R}^n} + \int_{[s + \beta, b]_\mathbb{T}} \chi_\Pi(\tau, \beta) \Delta\tau + L \int_{[s + \beta, t]_\mathbb{T}} \|\varepsilon_\Pi(\tau, \beta)\|_{\mathbb{R}^n} \Delta\tau,$$

for every $t \in [s + \beta, b]_\mathbb{T}$, and it follows from Lemma 2 that $\|\varepsilon_\Pi(t, \beta)\| \leq \Upsilon_\Pi(\beta) e_L(b, s)$, for every $t \in [s + \beta, b]_\mathbb{T}$, where $\Upsilon_\Pi(\beta) = \|\varepsilon_\Pi(s + \beta, \beta)\|_{\mathbb{R}^n} + \int_{[s + \beta, b]_\mathbb{T}} \chi_\Pi(\tau, \beta) \Delta\tau$.

To conclude, it remains to prove that $\Upsilon_\Pi(\beta)$ converges to 0 as β tends to 0. First, since $\theta_\Pi(\cdot, \beta)$ converges uniformly to $q(\cdot, u, q_a)$ on $[s + \beta, b]_\mathbb{T}$ as β tends to 0 and since $\partial f / \partial q$ is uniformly continuous on K , we infer that $\int_{[s + \beta, b]_\mathbb{T}} \chi_\Pi(\tau, \beta) \Delta\tau$ converges to 0 as β tends to 0. Second, let us prove that $\|\varepsilon_\Pi(s + \beta, \beta)\|_{\mathbb{R}^n}$ converges to 0 as β tends to 0. By continuity, $w_\Pi(s + \beta, u, q_a)$ converges to $w_\Pi(s, u, q_a)$ as β to 0. Moreover, since $q(\cdot, u_\Pi(\cdot, \beta), q_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_\mathbb{T}$ as β tends to 0 and since f is uniformly continuous on K , it follows that $f(q(\cdot, u_\Pi(\cdot, \beta), q_a), z, t)$ converges uniformly to $f(q(\cdot, u, q_a), z, t)$ on $[a, b]_\mathbb{T}$ as β tends to 0. Therefore, it suffices to note that

$$\frac{1}{\beta} \int_{[s, s + \beta]_\mathbb{T}} f(q(\tau, u, q_a), z, \tau) - f(q(\tau, u, q_a), u(\tau), \tau) \Delta\tau$$

converges to $w_\Pi(s, u, q_a) = f(q(s, u, q_a), z, s) - f(q(s, u, q_a), u(s), s)$ as β tends to 0 since s is a Δ -Lebesgue point of $f(q(\cdot, u, q_a), z, t)$ and of $f(q(\cdot, u, q_a), u, t)$. Then $\|\varepsilon_\Pi(s + \beta, \beta)\|$ converges to 0 as β tends to 0, and hence $\Upsilon_\Pi(\beta)$ converges to 0 as well. \square

Lemma 9. Let $R > \|u\|_{L^\infty_\mathbb{T}([a, b]_\mathbb{T}, \mathbb{R}^m)}$ and let $(u_k, q_{a,k})_{k \in \mathbb{N}}$ be a sequence of elements of $E(u, q_a, R)$. If u_k converges to u Δ -a.e. on $[a, b]_\mathbb{T}$ and $q_{a,k}$ converges to q_a as k tends to $+\infty$, then $w_\Pi(\cdot, u_k, q_{a,k})$ converges uniformly to $w_\Pi(\cdot, u, q_a)$ on $[s, b]_\mathbb{T}$ as k tends to $+\infty$.

Proof. The proof is similar to the one of Lemma 8, replacing $\sigma(r)$ with s . \square

3.2.4 Variation of the initial condition q_a

Let $q'_a \in \mathbb{R}^n$.

Lemma 10. *There exists $\gamma_0 > 0$ such that $(u, q_a + \gamma q'_a) \in \mathcal{UQ}_{\text{ad}}^b$ for every $\gamma \in [0, \gamma_0]$.*

Proof. Let $R = \|u\|_{L_T^\infty([a, b]_{\mathbb{T}}, \mathbb{R}^m)} + 1 > \|u\|_{L_T^\infty([a, b]_{\mathbb{T}}, \mathbb{R}^m)}$. We use the notations K , L , ν_R and η_R , associated with (u, q_a, R) , defined in Lemma 1 and in its proof.

There exists $\gamma_0 > 0$ such that $\|q_a + \gamma q'_a - q_a\|_{\mathbb{R}^n} = \gamma \|q'_a\|_{\mathbb{R}^n} \leq \eta_R$ for every $\gamma \in [0, \gamma_0]$, and hence $(u, q_a + \gamma q'_a) \in E(u, q_a, R)$. Then the claim follows from Lemma 1. \square

Lemma 11. *The mapping*

$$\begin{aligned} F_{(u, q_a, q'_a)} : ([0, \gamma_0], |\cdot|) &\longrightarrow (\mathcal{C}([a, b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_\infty) \\ \gamma &\longmapsto q(\cdot, u, q_a + \gamma q'_a) \end{aligned}$$

is Lipschitzian. In particular, for every $\gamma \in [0, \gamma_0]$, $q(\cdot, u, q_a + \gamma q'_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$ as γ tends to 0.

Proof. We use the notations of proof of Lemma 10. From Lemma 1, there exists $C \geq 0$ (Lipschitz constant of $F_{(u, q_a, R)}$) such that

$$\begin{aligned} \|q(\cdot, u, q_a + \gamma^2 q'_a) - q(\cdot, u, q_a + \gamma^1 q'_a)\|_\infty &\leq C d_{\mathcal{UQ}_{\text{ad}}^b}((u, q_a + \gamma^2 q'_a), (u, q_a + \gamma^1 q'_a)) \\ &= C |\gamma^2 - \gamma^1| \|q'_a\|_{\mathbb{R}^n}. \end{aligned}$$

for all γ^1 and γ^2 in $[0, \gamma_0]$. \square

According to [16, Theorem 3], we define the *variation vector* $w_{q'_a}(\cdot, u, q_a)$ associated with the perturbation q'_a as the unique solution on $[a, b]_{\mathbb{T}}$ of the linear Δ -Cauchy problem

$$w^\Delta(t) = \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t) w(t), \quad w(a) = q'_a. \quad (23)$$

Proposition 3. *The mapping*

$$\begin{aligned} F_{(u, q_a, q'_a)} : ([0, \gamma_0], |\cdot|) &\longrightarrow (\mathcal{C}([a, b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_\infty) \\ \gamma &\longmapsto q(\cdot, u, q_a + \gamma q'_a) \end{aligned} \quad (24)$$

is differentiable at 0, and one has $DF_{(u, q_a, q'_a)}(0) = w_{q'_a}(\cdot, u, q_a)$.

Proof. We use the notations of proof of Lemma 7. Note that, from Remark 14, $(q(t, u, q_a + \gamma q'_a), u(t), t) \in K$ for every $\gamma \in [0, \gamma_0]$ and for Δ -a.e. $t \in [a, b]_{\mathbb{T}}$. For every $\gamma \in]0, \gamma_0]$ and every $t \in [a, b]_{\mathbb{T}}$, we define

$$\varepsilon_{q'_a}(t, \gamma) = \frac{q(t, u, q_a + \gamma q'_a) - q(t, u, q_a)}{\gamma} - w_{q'_a}(t, u, q_a).$$

It suffices to prove that $\varepsilon_{q'_a}(\cdot, \gamma)$ converges uniformly to 0 on $[a, b]_{\mathbb{T}}$ as γ tends to 0. For every $\gamma \in]0, \gamma_0]$, the function $\varepsilon_{q'_a}(\cdot, \gamma)$ is absolutely continuous on $[a, b]_{\mathbb{T}}$, and $\varepsilon_{q'_a}(t, \gamma) = \varepsilon_{q'_a}(a, \gamma) + \int_{[a, t]_{\mathbb{T}}} \varepsilon_{q'_a}^\Delta(\tau, \gamma) \Delta\tau$, for every $t \in [a, b]_{\mathbb{T}}$, where

$$\varepsilon_{q'_a}^\Delta(t, \gamma) = \frac{f(q(t, u, q_a + \gamma q'_a), u(t), t) - f(q(t, u, q_a), u(t), t)}{\gamma} - \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t) w_{q'_a}(t, u, q_a),$$

for Δ -a.e. $t \in [a, b]_{\mathbb{T}}$. As in the proof of Proposition 1, it follows from the Mean Value Theorem that, for Δ -a.e. $t \in [a, b]_{\mathbb{T}}$, there exists $\theta_{q'_a}(t, \gamma) \in \mathbb{R}^n$, belonging to the segment of extremities $q(t, u, q_a)$ and $q(t, u, q_a + \gamma q'_a)$, such that

$$\begin{aligned} \varepsilon_{q'_a}^\Delta(t, \gamma) &= \frac{\partial f}{\partial q}(\theta_{q'_a}(t, \gamma), u(t), t) \varepsilon_{q'_a}(t, \gamma) \\ &\quad + \left(\frac{\partial f}{\partial q}(\theta_{q'_a}(t, \gamma), u(t), t) - \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t) \right) w_{q'_a}(t, u, q_a). \end{aligned}$$

Since $(\theta_{q'_a}(t, \gamma), u(t), t) \in K$ for Δ -a.e. $t \in [a, b]_{\mathbb{T}}$, it follows that

$$\|\varepsilon_{q'_a}^\Delta(t, \gamma)\|_{\mathbb{R}^n} \leq \chi_{q'_a}(t, \gamma) + L \|\varepsilon_{q'_a}(t, \gamma)\|_{\mathbb{R}^n},$$

where $\chi_{q'_a}(t, \gamma) = \left\| \left(\frac{\partial f}{\partial q}(\theta_{q'_a}(t, \gamma), u(t), t) - \frac{\partial f}{\partial q}(q(t, u, q_a), u(t), t) \right) \times w_{q'_a}(t, u, q_a) \right\|_{\mathbb{R}^n}$. Hence

$$\|\varepsilon_{q'_a}(t, \gamma)\|_{\mathbb{R}^n} \leq \|\varepsilon_{q'_a}(a, \gamma)\|_{\mathbb{R}^n} + \int_{[a, b]_{\mathbb{T}}} \chi_{q'_a}(\tau, \gamma) \Delta\tau + L \int_{[a, t]_{\mathbb{T}}} \|\varepsilon_{q'_a}(\tau, \gamma)\|_{\mathbb{R}^n} \Delta\tau,$$

for every $t \in [a, b]_{\mathbb{T}}$, and it follows from Lemma 2 that $\|\varepsilon_{q'_a}(t, \gamma)\| \leq \Upsilon_{q'_a}(\gamma) e_L(b, a)$, for every $t \in [a, b]_{\mathbb{T}}$, where $\Upsilon_{q'_a}(\gamma) = \|\varepsilon_{q'_a}(a, \gamma)\|_{\mathbb{R}^n} + \int_{[a, b]_{\mathbb{T}}} \chi_{q'_a}(\tau, \gamma) \Delta\tau$.

To conclude, it remains to prove that $\Upsilon_{q'_a}(\gamma)$ converges to 0 as γ tends to 0. First, since $\theta_{q'_a}(\cdot, \gamma)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$ as γ tends to 0 and since $\partial f / \partial q$ is uniformly continuous on K , we infer that $\int_{[a, b]_{\mathbb{T}}} \chi_{q'_a}(\tau, \gamma) \Delta\tau$ tends to 0 when $\gamma \rightarrow 0$. Second, it is easy to see that $\varepsilon_{q'_a}(a, \gamma) = 0$ for every $\gamma \in]0, \gamma_0]$. The conclusion follows. \square

Lemma 12. *Let $R > \|u\|_{L_{\mathbb{T}}^\infty([a, b]_{\mathbb{T}}, \mathbb{R}^m)}$ and let $(u_k, q_{a,k})_{k \in \mathbb{N}}$ be a sequence of elements of $E(u, q_a, R)$. If u_k converges to u Δ -a.e. on $[a, b]_{\mathbb{T}}$ and $q_{a,k}$ converges to q_a in \mathbb{R}^n as k tends to $+\infty$, then $w_{q'_a}(\cdot, u_k, q_{a,k})$ converges uniformly to $w_{q'_a}(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$ as k tends to $+\infty$.*

Proof. The proof is similar to the one of Lemma 8, replacing $\sigma(r)$ with a . \square

3.3 Proof of the PMP

Throughout this section we consider $(\mathbf{OCP})_{\mathbb{T}}$ with a fixed final time $b \in \mathbb{T} \setminus \{a\}$. We proceed as is very usual (see e.g. [40, 44]) by considering the *augmented control system* in \mathbb{R}^{n+1}

$$\bar{q}^\Delta(t) = \bar{f}(\bar{q}(t), u(t), t), \quad (25)$$

with $\bar{q} = (q, q^0)^T \in \mathbb{R}^n \times \mathbb{R}$, the augmented state, and $\bar{f} : \mathbb{R}^{n+1} \times \mathbb{R}^m \times \mathbb{T} \rightarrow \mathbb{R}^{n+1}$, the augmented dynamics, defined by $\bar{f}(\bar{q}, u, t) = (f(q, u, t), f^0(q, u, t))^T$. The additional coordinate q^0 stands for the cost, and we will always impose as an initial condition $q^0(a) = 0$, so that $q^0(b) = C(b, u) = \int_{[a, b]_{\mathbb{T}}} f^0(q(\tau), u(\tau), \tau) \Delta\tau$. The function $\bar{g} : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \rightarrow \mathbb{R}^j$ is defined by $\bar{g}(\bar{q}_1, \bar{q}_2) = g(q_1, q_2)$, where $\bar{q}_i = (q_i, q_i^0)$ for $i = 1, 2$. Note that \bar{f} does not depend on q^0 and that \bar{g} does not depend on q_1^0 nor on q_2^0 . Note as well that the Hamiltonian of $(\mathbf{OCP})_{\mathbb{T}}$ is written as $H(q, u, p, p^0, t) = \langle \bar{p}, \bar{f}(\bar{q}, u, t) \rangle_{\mathbb{R}^{n+1}}$.

With these notations, $(\mathbf{OCP})_{\mathbb{T}}$ consists of determining a trajectory $\bar{q}^*(\cdot) = (q^*(\cdot), q^{0*}(\cdot))$ defined on $[a, b]_{\mathbb{T}}$, solution of (25) and associated with a control $u^* \in L_{\mathbb{T}}^\infty([a, b]_{\mathbb{T}}; \Omega)$, minimizing $q^0(b)$ over all possible trajectories $\bar{q}(\cdot) = (q(\cdot), q^0(\cdot))$ defined on $[a, b]_{\mathbb{T}}$, solutions of (25) and associated with an admissible control $u \in L_{\mathbb{T}}^\infty([a, b]_{\mathbb{T}}; \Omega)$ and satisfying $\bar{g}(\bar{q}(a), \bar{q}(b)) \in S$.

In what follows, let $\bar{q}^*(\cdot)$ be such an optimal trajectory. Set $q_a^* = q^*(a)$. We are going to apply first Ekeland's Variational Principle to a well chosen functional in an appropriate complete metric space, and then, using needle-like variations as defined previously (applied to the augmented system, that is, with the dynamics \bar{f}), we are going to derive some inequalities, finally resulting into the desired statement of the PMP.

3.3.1 Application of Ekeland's Variational Principle

For the completeness, we recall Ekeland's Variational Principle.

Theorem 2 ([23]). *Let (E, d_E) be a complete metric space and $J : E \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous function which is bounded below. Let $\varepsilon > 0$ and $u^* \in E$ such that $J(u^*) \leq \inf_{u \in E} J(u) + \varepsilon$. Then there exists $u_\varepsilon \in E$ such that $d_E(u_\varepsilon, u^*) \leq \sqrt{\varepsilon}$ and $J(u_\varepsilon) \leq J(u) + \sqrt{\varepsilon} d_E(u, u_\varepsilon)$ for every $u \in E$.*

Recall from Lemma 1 that, for $R > \|u^*\|_{L^\infty_T([a, b[\mathbb{T}], \mathbb{R}^m)}$, the set $E(u^*, \bar{q}_a^*, R)$ defined in this lemma is contained in $\mathcal{UQ}_{\text{ad}}^b$. To take into account the set Ω of constraints on the controls, we define

$$E_\Omega^R = \{(u, \bar{q}_a) \in \mathcal{U} \times \mathbb{R}^{n+1} \mid \bar{q}_a = (q_a, 0), (u, q_a) \in E(u^*, \bar{q}_a^*, R), u \in L^\infty_T([a, b[\mathbb{T}]; \Omega)\}.$$

Using the fact that Ω is closed⁴, it clearly follows from the (partial) converse of Lebesgue's Dominated Convergence Theorem that $(E_\Omega^R, d_{\mathcal{UQ}_{\text{ad}}^b})$ is a complete metric space.

Before applying Ekeland's Variational Principle in this space, let us introduce several notations and recall basic facts in order to handle the convex set S . We denote by d_S the distance function to S defined by $d_S(x) = \inf_{x' \in S} \|x - x'\|_{\mathbb{R}^j}$, for every $x \in \mathbb{R}^j$. Recall that, for every $x \in \mathbb{R}^j$, there exists a unique element $P_S(x) \in S$ (projection of x onto S) such that $d_S(x) = \|x - P_S(x)\|_{\mathbb{R}^j}$. It is characterized by the property $\langle x - P_S(x), x' - P_S(x) \rangle_{\mathbb{R}^j} \leq 0$ for every $x' \in S$. Moreover, the projection mapping P_S is 1-Lipschitz continuous. Furthermore, there holds $x - P_S(x) \in \mathcal{O}_S(P_S(x))$ for every $x \in \mathbb{R}^j$ (where $\mathcal{O}_S(x)$ is defined by (6)). We recall the following obvious lemmas.

Lemma 13. *Let $(x_k)_{k \in \mathbb{N}}$ be a sequence of points of \mathbb{R}^j and $(\zeta_k)_{k \in \mathbb{N}}$ be a sequence of nonnegative real numbers such that $x_k \rightarrow x \in S$ and $\zeta_k(x_k - P_S(x_k)) \rightarrow x' \in \mathbb{R}^j$ as $k \rightarrow +\infty$. Then $x' \in \mathcal{O}_S(x)$.*

Lemma 14. *The function $x \mapsto d_S^2(x)^2$ is differentiable on \mathbb{R}^j , and $dd_S^2(x) \cdot x' = 2\langle x - P_S(x), x' \rangle_{\mathbb{R}^j}$.*

We are now in a position to apply Ekeland's Variational Principle. For every $\varepsilon > 0$ such that $\sqrt{\varepsilon} < \min(\nu_R, \eta_R)$, we consider the functional $J_\varepsilon^R : (E_\Omega^R, d_{\mathcal{UQ}_{\text{ad}}^b}) \rightarrow \mathbb{R}^+$ defined by

$$J_\varepsilon^R(u, \bar{q}_a) = (\max(q^0(b, u, \bar{q}_a) - q^{0*}(b) + \varepsilon, 0)^2 + d_S^2(\bar{g}(\bar{q}_a, \bar{q}(b, u, \bar{q}_a))))^{1/2}.$$

Since $F_{(u^*, \bar{q}_a^*, R)}$ (by Lemma 3), \bar{g} and d_S are continuous, it follows that J_ε^R is continuous on $(E_\Omega^R, d_{\mathcal{UQ}_{\text{ad}}^b})$. Moreover, one has $J_\varepsilon^R(u^*, \bar{q}_a^*) = \varepsilon$ and $J_\varepsilon^R(u, \bar{q}_a) > 0$ for every $(u, \bar{q}_a) \in E_\Omega^R$. It follows from Ekeland's Variational Principle that, for every $\varepsilon > 0$ such that $\sqrt{\varepsilon} < \min(\nu_R, \eta_R)$, there exists $(u_\varepsilon^R, \bar{q}_{a, \varepsilon}^R) \in E_\Omega^R$ such that $d_{\mathcal{UQ}_{\text{ad}}^b}((u_\varepsilon^R, \bar{q}_{a, \varepsilon}^R), (u^*, \bar{q}_a^*)) \leq \sqrt{\varepsilon}$ and

$$-\sqrt{\varepsilon} d_{\mathcal{UQ}_{\text{ad}}^b}((u, \bar{q}_a), (u_\varepsilon^R, \bar{q}_{a, \varepsilon}^R)) \leq J_\varepsilon^R(u, \bar{q}_a) - J_\varepsilon^R(u_\varepsilon^R, \bar{q}_{a, \varepsilon}^R), \quad (26)$$

for every $(u, \bar{q}_a) \in E_\Omega^R$. In particular, u_ε^R converges to u^* in $L^1_T([a, b[\mathbb{T}], \mathbb{R}^m)$ and $\bar{q}_{a, \varepsilon}^R$ converges to \bar{q}_a^* as ε tends to 0. Besides, setting

$$\psi_\varepsilon^{0R} = \frac{-1}{J_\varepsilon^R(u_\varepsilon^R, \bar{q}_{a, \varepsilon}^R)} \max(q^0(b, u_\varepsilon^R, \bar{q}_{a, \varepsilon}^R) - q^{0*}(b) + \varepsilon, 0) \leq 0 \quad (27)$$

⁴Note that the assumption Ω closed is used (only) here in a crucial way. In the proof of the classical continuous-time PMP this assumption is not required because the Ekeland distance which is then used is defined by $\rho(u, v) = \mu_L(\{t \in [a, b] \mid u(t) \neq v(t)\})$, and obviously the set of measurable functions $u : [a, b] \rightarrow \Omega$ endowed with this distance is complete, under the sole assumption that Ω is measurable. In the discrete-time setting and a fortiori in the general time scale setting, this distance cannot be used any more. Here we use the distance $d_{\mathcal{UQ}_{\text{ad}}^b}$ defined by (17) but then to ensure completeness it is required to assume that Ω is closed.

and

$$\psi_\varepsilon^R = \frac{-1}{J_\varepsilon^R(u_\varepsilon^R, \bar{q}_{a,\varepsilon}^R)} (\bar{g}(\bar{q}_{a,\varepsilon}^R, \bar{q}(b, u_\varepsilon^R, \bar{q}_{a,\varepsilon}^R)) - \text{Ps}(\bar{g}(\bar{q}_{a,\varepsilon}^R, \bar{q}(b, u_\varepsilon^R, \bar{q}_{a,\varepsilon}^R)))) \in \mathbb{R}^j, \quad (28)$$

note that $|\psi_\varepsilon^{0R}|^2 + \|\psi_\varepsilon^R\|_{\mathbb{R}^j}^2 = 1$ and $-\psi_\varepsilon^R \in \mathcal{O}_S(\text{Ps}(\bar{g}(\bar{q}_{a,\varepsilon}^R, \bar{q}(b, u_\varepsilon^R, \bar{q}_{a,\varepsilon}^R))))$.

Using a compactness argument, the continuity of $F_{(u^*, \bar{q}_a^*, R)}$ and the \mathcal{C}^1 regularity of \bar{g} , and the (partial) converse of the Dominated Convergence Theorem, we infer that there exists a sequence $(\varepsilon_k)_{k \in \mathbb{N}}$ of positive real numbers converging to 0 such that $u_{\varepsilon_k}^R$ converges to u^* Δ -a.e. on $[a, b]_{\mathbb{T}}$, $\bar{q}_{a,\varepsilon_k}^R$ converges to \bar{q}_a^* , $\bar{g}(\bar{q}_{a,\varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R))$ converges to $\bar{g}(\bar{q}_a^*, \bar{q}^*(b)) \in S$, $d\bar{g}(\bar{q}_{a,\varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R))$ converges to $d\bar{g}(\bar{q}_a^*, \bar{q}^*(b))$, $\psi_{\varepsilon_k}^{0R}$ converges to some $\psi^{0R} \leq 0$, and $\psi_{\varepsilon_k}^R$ converges to some $\psi^R \in \mathbb{R}^j$ as k tends to $+\infty$, with $|\psi^{0R}|^2 + \|\psi^R\|_{\mathbb{R}^j}^2 = 1$ and $-\psi^R \in \mathcal{O}_S(\bar{g}(\bar{q}_a^*, \bar{q}^*(b)))$ (see Lemma 13).

In the three next lemmas, we use the inequality (26) respectively with needle-like variations of $u_{\varepsilon_k}^R$ at right-scattered points and then at right-dense points, and variations of $\bar{q}_{a,\varepsilon_k}^R$, and infer some crucial inequalities by taking the limit in k . Note that these variations were defined in Section 3.2 for any dynamics f , and that we apply them here to the augmented system (25), associated with the augmented dynamics \bar{f} .

Lemma 15. *For every $r \in [a, b]_{\mathbb{T}} \cap \text{RS}$ and every $y \in \mathcal{D}_{\text{stab}}^\Omega(u^*(r))$, considering the needle-like variation $\Pi = (r, y)$ at the right-scattered point r as defined in Section 3.2.2, there holds*

$$\psi^{0R} w_\Pi^0(b, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b)) \right)^T \psi^R, w_\Pi(b, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n} \leq 0, \quad (29)$$

where the variation vector $\bar{w}_\Pi = (w_\Pi, w_\Pi^0)$ is defined by (19) (replacing f with \bar{f}).

Proof. Since $u_{\varepsilon_k}^R$ converges to u^* Δ -a.e. on $[a, b]_{\mathbb{T}}$, it follows that $u_{\varepsilon_k}^R(r)$ converges to $u^*(r)$ as k tends to $+\infty$. Hence $y \in \mathcal{D}^\Omega(u_{\varepsilon_k}^R(r))$ and $\|u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m} < R$ for k sufficiently large. Fixing such a large integer k , we recall that $u_{\varepsilon_k, \Pi}^R(\cdot, \alpha) \in L_{\mathbb{T}}^\infty([a, b]_{\mathbb{T}}; \Omega)$ for every $\alpha \in \mathcal{D}^\Omega(u_{\varepsilon_k}^R(r), y)$, and

$$\begin{aligned} \|u_{\varepsilon_k, \Pi}^R(\cdot, \alpha)\|_{L_{\mathbb{T}}^\infty([a, b]_{\mathbb{T}}; \mathbb{R}^m)} &\leq \max(\|u_{\varepsilon_k}^R\|_{L_{\mathbb{T}}^\infty([a, b]_{\mathbb{T}}; \mathbb{R}^m)}, \|u_{\varepsilon_k, \Pi}^R(r, \alpha)\|_{\mathbb{R}^m}) \\ &\leq \max(R, \|u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m} + \alpha\|y - u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m}), \end{aligned}$$

and

$$\begin{aligned} \|u_{\varepsilon_k, \Pi}^R(\cdot, \alpha) - u^*\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}; \mathbb{R}^m)} &\leq \|u_{\varepsilon_k, \Pi}^R(\cdot, \alpha) - u_{\varepsilon_k}^R\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}; \mathbb{R}^m)} + \|u_{\varepsilon_k}^R - u^*\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}; \mathbb{R}^m)} \\ &\leq \alpha\mu(r)\|y - u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m} + \sqrt{\varepsilon_k}. \end{aligned}$$

Therefore $(u_{\varepsilon_k, \Pi}^R(\cdot, \alpha), \bar{q}_{a,\varepsilon_k}^R) \in E_\Omega^R$ for every $\alpha \in \mathcal{D}^\Omega(u_{\varepsilon_k}^R(r), y)$ sufficiently small. It then follows from (26) that

$$-\sqrt{\varepsilon_k} \|u_{\varepsilon_k, \Pi}^R(\cdot, \alpha) - u_{\varepsilon_k}^R\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}; \mathbb{R}^m)} \leq J_k^R(u_{\varepsilon_k, \Pi}^R(\cdot, \alpha), \bar{q}_{a,\varepsilon_k}^R) - J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R),$$

and thus

$$-\sqrt{\varepsilon_k} \mu(r) \|y - u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m} \leq \frac{J_k^R(u_{\varepsilon_k, \Pi}^R(\cdot, \alpha), \bar{q}_{a,\varepsilon_k}^R)^2 - J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)^2}{\alpha(J_k^R(u_{\varepsilon_k, \Pi}^R(\cdot, \alpha), \bar{q}_{a,\varepsilon_k}^R) + J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R))}.$$

Using Proposition 1, since \bar{g} does not depend on q_2^0 , we infer that

$$\begin{aligned} &\lim_{\alpha \rightarrow 0} \frac{J_k^R(u_{\varepsilon_k, \Pi}^R(\cdot, \alpha), \bar{q}_{a,\varepsilon_k}^R)^2 - J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)^2}{\alpha} \\ &= 2 \max(q^0(b, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R) - q^{0*}(b) + \varepsilon_k, 0) w_\Pi^0(b, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R) \\ &\quad + 2 \left\langle \bar{g}(\bar{q}_{a,\varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)) - \text{Ps}(\bar{g}(\bar{q}_{a,\varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R))) \right. \\ &\quad \left. \frac{\partial \bar{g}}{\partial q_2}(\bar{q}_{a,\varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)) w_\Pi(b, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R) \right\rangle_{\mathbb{R}^j}. \end{aligned}$$

Since $J_k^R(u_{\varepsilon_k, \Pi}^R(\cdot, \alpha), \bar{q}_{a, \varepsilon_k}^R)$ converges to $J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)$ as α tends to 0, using (27) and (28) it follows that

$$-\sqrt{\varepsilon_k} \mu(r) \|y - u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m} \leq -\psi_{\varepsilon_k}^{0R} w_{\Pi}^0(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) - \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)) \right)^T \psi_{\varepsilon_k}^R, w_{\Pi}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) \right\rangle_{\mathbb{R}^n}.$$

By letting k tend to $+\infty$, and using Lemma 6, the lemma follows. \square

Denote by $\mathcal{L}_{[a, b]_{\mathbb{T}}}^R$ the set of (Lebesgue) times $t \in [a, b]_{\mathbb{T}}$ such that $t \in \mathcal{L}_{[a, b]_{\mathbb{T}}}(f(q(\cdot, u^*, q_a^*), u^*, t))$, such that $t \in \mathcal{L}_{[a, b]_{\mathbb{T}}}(f(q(\cdot, u_{\varepsilon_k}^R, q_{a, k}^R), u_{\varepsilon_k}^R, t))$ for every $k \in \mathbb{N}$, and such that $u_{\varepsilon_k}^R(t)$ converges to $u^*(t)$ as k tends to $+\infty$. There holds $\mu_{\Delta}(\mathcal{L}_{[a, b]_{\mathbb{T}}}^R) = \mu_{\Delta}([a, b]_{\mathbb{T}}) = b - a$.

Lemma 16. *For every $s \in \mathcal{L}_{[a, b]_{\mathbb{T}}}^R \cap \text{RD}$ and any $z \in \Omega \cap \overline{B}_{\mathbb{R}^m}(0, R)$, considering the needle-like variation $\Pi = (s, z)$ as defined in Section 3.2.3, there holds*

$$\psi^{0R} w_{\Pi}^0(b, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b)) \right)^T \psi^R, w_{\Pi}(b, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n} \leq 0, \quad (30)$$

where the variation vector $\bar{w}_{\Pi} = (w_{\Pi}, w_{\Pi}^0)$ is defined by (21) (replacing f with \bar{f}).

Proof. For every $k \in \mathbb{N}$ and any $\beta \in \mathcal{V}_s^b$, we recall that $u_{\varepsilon_k, \Pi}^R(\cdot, \beta) \in L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}, \Omega)$ and

$$\|u_{\varepsilon_k, \Pi}^R(\cdot, \beta)\|_{L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}, \mathbb{R}^m)} \leq \max(\|u_{\varepsilon_k}^R\|_{L_{\mathbb{T}}^{\infty}([a, b]_{\mathbb{T}}, \mathbb{R}^m)}, \|z\|_{\mathbb{R}^m}) \leq R,$$

and

$$\begin{aligned} \|u_{\varepsilon_k, \Pi}^R(\cdot, \beta) - u^*\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}, \mathbb{R}^m)} &\leq \|u_{\varepsilon_k, \Pi}^R(\cdot, \beta) - u_{\varepsilon_k}^R\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}, \mathbb{R}^m)} + \|u_{\varepsilon_k}^R - u^*\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}, \mathbb{R}^m)} \\ &\leq 2R\beta + \sqrt{\varepsilon_k}. \end{aligned}$$

Therefore $(u_{\varepsilon_k, \Pi}^R(\cdot, \beta), \bar{q}_{a, \varepsilon_k}^R) \in E_{\Omega}^R$ for $\beta \in \mathcal{V}_s^b$ sufficiently small. It then follows from (26) that

$$-\sqrt{\varepsilon_k} \|u_{\varepsilon_k, \Pi}^R(\cdot, \beta) - u_{\varepsilon_k}^R\|_{L_{\mathbb{T}}^1([a, b]_{\mathbb{T}}, \mathbb{R}^m)} \leq J_k^R(u_{\varepsilon_k, \Pi}^R(\cdot, \beta), \bar{q}_{a, \varepsilon_k}^R) - J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R),$$

and thus

$$-2R\sqrt{\varepsilon_k} \leq \frac{J_k^R(u_{\varepsilon_k, \Pi}^R(\cdot, \beta), \bar{q}_{a, \varepsilon_k}^R)^2 - J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)^2}{\beta(J_k^R(u_{\varepsilon_k, \Pi}^R(\cdot, \beta), \bar{q}_{a, \varepsilon_k}^R) + J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R))}.$$

Using Proposition (2), since \bar{g} does not depend on q_2^0 , we infer that

$$\begin{aligned} &\lim_{\beta \rightarrow 0} \frac{J_k^R(u_{\varepsilon_k, \Pi}^R(\cdot, \beta), \bar{q}_{a, \varepsilon_k}^R)^2 - J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)^2}{\beta} \\ &= 2 \max(q^0(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) - q^{0*}(b) + \varepsilon_k, 0) w_{\Pi}^0(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) \\ &\quad + 2 \left\langle \bar{g}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)) - \text{Ps}(\bar{g}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R))), \right. \\ &\quad \left. \frac{\partial \bar{g}}{\partial q_2}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)) w_{\Pi}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) \right\rangle_{\mathbb{R}^j}. \end{aligned}$$

Since $J_k^R(u_{\varepsilon_k, \Pi}^R(\cdot, \beta), \bar{q}_{a, \varepsilon_k}^R)$ converges to $J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)$ as α tends to 0, using (27) and (28) it follows that

$$-2R\sqrt{\varepsilon_k} \leq -\psi_{\varepsilon_k}^{0R} w_{\Pi}^0(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) - \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)) \right)^T \psi_{\varepsilon_k}^R, w_{\Pi}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) \right\rangle_{\mathbb{R}^n}.$$

By letting k tend to $+\infty$, and using Lemma 9, the lemma follows. \square

Lemma 17. For every $\bar{q}_a \in \mathbb{R}^n \times \{0\}$, considering the variation of initial point as defined in Section 3.2.4, there holds

$$\psi^{0R} w_{\bar{q}_a}^0(b, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b)) \right)^\top \psi^R, w_{\bar{q}_a}(b, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n} \leq - \left\langle \left(\frac{\partial \bar{g}}{\partial q_1}(\bar{q}_a^*, \bar{q}^*(b)) \right)^\top \psi^R, q_a \right\rangle_{\mathbb{R}^n}, \quad (31)$$

where the variation vector $\bar{w}_{\bar{q}_a} = (w_{\bar{q}_a}, w_{\bar{q}_a}^0)$ is defined by (23) (replacing f with \bar{f}).

Proof. For every $k \in \mathbb{N}$ and every $\gamma \geq 0$, one has

$$\|\bar{q}_{a, \varepsilon_k}^R + \gamma \bar{q}_a - \bar{q}_a^*\|_{\mathbb{R}^n} \leq \gamma \|\bar{q}_a\|_{\mathbb{R}^n} + \|\bar{q}_{a, \varepsilon_k}^R - \bar{q}_a^*\|_{\mathbb{R}^n} \leq \gamma \|\bar{q}_a\|_{\mathbb{R}^n} + \sqrt{\varepsilon_k}.$$

Therefore $(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R + \gamma \bar{q}_a) \in E_\Omega^R$ for $\gamma \geq 0$ sufficiently small. It then follows from (26) that

$$-\sqrt{\varepsilon_k} \|\bar{q}_{a, \varepsilon_k}^R + \gamma \bar{q}_a - \bar{q}_{a, \varepsilon_k}^R\|_{\mathbb{R}^n} \leq J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R + \gamma \bar{q}_a) - J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R),$$

and thus

$$-\sqrt{\varepsilon_k} \|\bar{q}_a\|_{\mathbb{R}^n} \leq \frac{J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R + \gamma \bar{q}_a)^2 - J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)^2}{\gamma (J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R + \gamma \bar{q}_a) + J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R))}.$$

Using Proposition (3), since \bar{g} does not depend on q_1^0 and q_2^0 , we infer that

$$\begin{aligned} & \lim_{\gamma \rightarrow 0} \frac{J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R + \gamma \bar{q}_a)^2 - J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)^2}{\gamma} \\ &= 2 \max(q^0(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) - q^{0*}(b) + \varepsilon_k, 0) w_{\bar{q}_a}^0(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) \\ & \quad + 2 \left\langle \bar{g}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)) - \text{Ps}(\bar{g}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R))), \frac{\partial \bar{g}}{\partial q_1}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)) q_a \right\rangle_{\mathbb{R}^j} \\ & \quad + 2 \left\langle \bar{g}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)) - \text{Ps}(\bar{g}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R))), \right. \\ & \quad \left. \frac{\partial \bar{g}}{\partial q_2}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)) w_{\bar{q}_a}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) \right\rangle_{\mathbb{R}^j}. \end{aligned}$$

Since $J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R + \gamma \bar{q}_a)$ converges to $J_k^R(u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)$ as γ tends to 0, using (27) and (28) it follows that

$$\begin{aligned} -\sqrt{\varepsilon_k} \|\bar{q}_a\| &\leq -\psi_{\varepsilon_k}^{0R} w_{\bar{q}_a}^0(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) - \left\langle \left(\frac{\partial \bar{g}}{\partial q_1}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)) \right)^\top \psi_{\varepsilon_k}^R, q_a \right\rangle_{\mathbb{R}^n} \\ &\quad - \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_{a, \varepsilon_k}^R, \bar{q}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R)) \right)^\top \psi_{\varepsilon_k}^R, w_{\bar{q}_a}(b, u_{\varepsilon_k}^R, \bar{q}_{a, \varepsilon_k}^R) \right\rangle_{\mathbb{R}^n}. \end{aligned}$$

By letting k tend to $+\infty$, and using Lemma 12, the lemma follows. \square

At this step, we have obtained in the three previous lemmas the three fundamental inequalities (29), (30) and (31), valuable for any $R > \|u^*\|_{L_T^\infty([a, b]_{\mathbb{T}}, \mathbb{R}^m)}$. Recall that $|\psi^{0R}|^2 + \|\psi^R\|_{\mathbb{R}^j}^2 = 1$ and $-\psi^R \in \mathcal{O}_S(\bar{g}(\bar{q}_a^*, \bar{q}^*(b)))$. Then, considering a sequence of real numbers R_ℓ converging to $+\infty$ as ℓ tends to $+\infty$, we infer that there exist $\psi^0 \leq 0$ and $\psi \in \mathbb{R}^j$ such that ψ^{0R_ℓ} converges to ψ^0 and ψ^{R_ℓ} converges to ψ as ℓ tends to $+\infty$, and moreover $|\psi^0|^2 + \|\psi\|_{\mathbb{R}^j}^2 = 1$ and $-\psi \in \mathcal{O}_S(\bar{g}(\bar{q}_a^*, \bar{q}^*(b)))$ (since $\mathcal{O}_S(\bar{g}(\bar{q}_a^*, \bar{q}^*(b)))$ is a closed subset of \mathbb{R}^j).

We set $\mathcal{L}_{[a, b]_{\mathbb{T}}} = \bigcap_{\ell \in \mathbb{N}} \mathcal{L}_{[a, b]_{\mathbb{T}}}^{R_\ell}$. Note that $\mu_\Delta(\mathcal{L}_{[a, b]_{\mathbb{T}}}) = \mu_\Delta([a, b]_{\mathbb{T}}) = b - a$. Taking the limit in ℓ in (29), (30) and (31), we get the following lemma.

Lemma 18. *We have the following variational inequalities.*

For every $r \in [a, b]_{\mathbb{T}} \cap \text{RS}$, and every $y \in \mathcal{D}_{\text{stab}}^{\Omega}(u^(r))$, there holds*

$$\psi^0 w_{\Pi}^0(b, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b)) \right)^{\text{T}} \psi, w_{\Pi}(b, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n} \leq 0, \quad (32)$$

where the variation vector $\bar{w}_{\Pi} = (w_{\Pi}, w_{\Pi}^0)$ associated with the needle-like variation $\Pi = (r, y)$ is defined by (19) (replacing f with \bar{f});

For every $s \in \mathcal{L}_{[a, b]_{\mathbb{T}}} \cap \text{RD}$, and every $z \in \Omega$, there holds

$$\psi^0 w_{\Pi}^0(b, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b)) \right)^{\text{T}} \psi, w_{\Pi}(b, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n} \leq 0, \quad (33)$$

where the variation vector $\bar{w}_{\Pi} = (w_{\Pi}, w_{\Pi}^0)$ associated with the needle-like variation $\Pi = (s, z)$ is defined by (21) (replacing f with \bar{f});

For every $\bar{q}_a \in \mathbb{R}^n \times \{0\}$, there holds

$$\psi^0 w_{\bar{q}_a}^0(b, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b)) \right)^{\text{T}} \psi, w_{\bar{q}_a}(b, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n} \leq - \left\langle \left(\frac{\partial \bar{g}}{\partial q_1}(\bar{q}_a^*, \bar{q}^*(b)) \right)^{\text{T}} \psi, q_a \right\rangle_{\mathbb{R}^n}, \quad (34)$$

where the variation vector $\bar{w}_{\bar{q}_a} = (w_{\bar{q}_a}, w_{\bar{q}_a}^0)$ associated with the variation \bar{q}_a of the initial point q_a^ is defined by (23) (replacing f with \bar{f}).*

This result concludes the application of Ekeland's Variational Principle. The last step of the proof consists of deriving the PMP from these inequalities.

3.3.2 End of the proof

We define $\bar{p}(\cdot) = (p(\cdot), p^0(\cdot))$ as the unique solution on $[a, b]_{\mathbb{T}}$ of the backward shifted linear Δ -Cauchy problem

$$\bar{p}^{\Delta}(t) = - \left(\frac{\partial \bar{f}}{\partial \bar{q}}(\bar{q}^*(t), u^*(t), t) \right)^{\text{T}} \bar{p}^{\sigma}(t), \quad \bar{p}(b) = \left(\left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b)) \right)^{\text{T}} \psi, \psi^0 \right)^{\text{T}}.$$

The existence and uniqueness of $\bar{p}(\cdot)$ are ensured by [16, Theorem 6]. Since \bar{f} does not depend on q^0 , it is clear that $p^0(\cdot)$ is constant, still denoted by p^0 (with $p^0 = \psi^0$).

Right-scattered points. Let $r \in [a, b]_{\mathbb{T}} \cap \text{RS}$ and $y \in \mathcal{D}_{\text{stab}}^{\Omega}(u^*(r))$. Since the function $t \mapsto \langle \bar{w}_{\Pi}(t, u^*, \bar{q}_a^*), \bar{p}(t) \rangle_{\mathbb{R}^{n+1}}$ is absolutely continuous, it holds $\langle \bar{p}(\cdot), \bar{w}_{\Pi}(\cdot, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}}^{\Delta} = 0$ Δ -almost everywhere on $[\sigma(r), b]_{\mathbb{T}}$ from the Leibniz formula (2) and hence the function $\langle \bar{p}(\cdot), \bar{w}_{\Pi}(\cdot, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}}$ is constant on $[\sigma(r), b]_{\mathbb{T}}$. It thus follows from (32) that

$$\begin{aligned} \langle \bar{p}(\sigma(r)), \bar{w}_{\Pi}(\sigma(r), u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}} &= \langle \bar{p}(b), \bar{w}_{\Pi}(b, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}} \\ &= \psi^0 w_{\Pi}^0(b, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b)) \right)^{\text{T}} \psi, w_{\Pi}(b, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n} \leq 0, \end{aligned}$$

and since $\bar{w}_{\Pi}(\sigma(r), u^*, \bar{q}_a^*) = \mu(r) \frac{\partial \bar{f}}{\partial u}(\bar{q}^*(r), u^*(r), r)(y - u^*(r))$, we finally get

$$\left\langle \frac{\partial H}{\partial u}(\bar{q}^*(r), u^*(r), \bar{p}^{\sigma}(r), r), y - u^*(r) \right\rangle_{\mathbb{R}^m} \leq 0.$$

Since this inequality holds for every $y \in \mathcal{D}_{\text{stab}}^{\Omega}(u^*(r))$, we easily prove that it holds as well for every $v \in \overline{\text{Co}}(\mathcal{D}_{\text{stab}}^{\Omega}(u^*(r)))$. This proves (8).

Right-dense points. Let $s \in \mathcal{L}_{[a, b]_{\mathbb{T}}} \cap \text{RD}$ and $z \in \Omega$. Since $t \mapsto \langle \bar{w}_{\Pi}(t, u^*, \bar{q}_a^*), \bar{p}(t) \rangle_{\mathbb{R}^{n+1}}$ is an absolutely continuous function, the Leibniz formula (2) yields $\langle \bar{p}(\cdot), \bar{w}_{\Pi}(\cdot, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}}^{\Delta} = 0$ Δ -almost everywhere on $[s, b]_{\mathbb{T}}$, and hence this function is constant on $[s, b]_{\mathbb{T}}$. It thus follows from (33) that

$$\begin{aligned} \langle \bar{p}(s), \bar{w}_{\Pi}(s, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}} &= \langle \bar{p}(b), \bar{w}_{\Pi}(b, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}} \\ &= \psi^0 w_{\Pi}^0(b, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b)) \right)^{\text{T}} \psi, w_{\Pi}(b, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n}, \end{aligned}$$

and since $\bar{w}_{\Pi}(s, u^*, \bar{q}_a^*) = \bar{f}(\bar{q}^*(s), z, s) - \bar{f}(\bar{q}(s), u^*(s), s)$, we finally get

$$\langle \bar{p}(s), \bar{f}(\bar{q}^*(s), z, s) \rangle_{\mathbb{R}^{n+1}} \leq \langle \bar{p}(s), \bar{f}(\bar{q}(s), u^*(s), s) \rangle_{\mathbb{R}^{n+1}}.$$

Since this inequality holds for every $z \in \Omega$, the maximization condition (9) follows.

Transversality conditions. The transversality condition on the adjoint vector p at the final time b has been obtained by definition (note that $-\psi \in \mathcal{O}_S(\bar{g}(\bar{q}_a^*, \bar{q}^*(b)))$ as mentioned previously). Let us now establish the transversality condition at the initial time a (left-hand equality of (10)). Let $\bar{q}_a \in \mathbb{R}^n \times \{0\}$. With the same arguments as before, we prove that the function $t \mapsto \langle \bar{w}_{\bar{q}_a}(t, u^*, \bar{q}_a^*), \bar{p}(t) \rangle_{\mathbb{R}^{n+1}}$ is constant on $[a, b]_{\mathbb{T}}$. It thus follows from (34) that

$$\begin{aligned} \langle \bar{p}(a), \bar{w}_{\bar{q}_a}(a, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}} &= \langle \bar{p}(b), \bar{w}_{\bar{q}_a}(b, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}} \\ &= \psi^0 w_{\bar{q}_a}^0(b, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial \bar{g}}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b)) \right)^{\text{T}} \psi, w_{\bar{q}_a}(b, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n} \\ &\leq - \left\langle \left(\frac{\partial \bar{g}}{\partial q_1}(\bar{q}_a^*, \bar{q}^*(b)) \right)^{\text{T}} \psi, q_a \right\rangle_{\mathbb{R}^n}, \end{aligned}$$

and since $\bar{w}_{\bar{q}_a}(a, u^*, \bar{q}_a^*) = \bar{q}_a = (q_a, 0)$, we finally get

$$\left\langle p(a, u^*, \bar{q}_a^*) + \left(\frac{\partial \bar{g}}{\partial q_1}(\bar{q}_a^*, \bar{q}^*(b)) \right)^{\text{T}} \psi, q_a \right\rangle_{\mathbb{R}^n} \leq 0.$$

Since this inequality holds for every $\bar{q}_a \in \mathbb{R}^n \times \{0\}$, the left-hand equality of (10) follows.

Free final time. Assume that the final time is not fixed in $(\text{OCP})_{\mathbb{T}}$, and let b^* be the final time associated with the optimal trajectory $q^*(\cdot)$. We assume moreover that b^* belongs to the interior of \mathbb{T} for the topology of \mathbb{R} . The proof of (11) then goes exactly as in the classical continuous-time case, and thus we do not provide any details. It suffices to consider variations of the final time b in a neighbourhood of b^* , and to modify accordingly the functional of Section 3.3.1 to which Ekeland's Variational Principle is applied.

To derive (12), we consider the change of variable $\tilde{t} = (t - a)/(b - a)$. The crucial remark is that, since it is an *affine* change of variable, Δ -derivatives of compositions work in the time scale setting as in the time-continuous case. Then it suffices to consider the resulting optimal control problem as a parametrized one with parameter b lying in a neighbourhood of b^* . Then (12) follows from the additional condition (14) of the PMP with parameters (see Remark 5), which is established hereafter.

PMP with parameters (Remark 5). To obtain the statement it suffices to apply the PMP to the control system associated with the dynamics defined by $\tilde{f}(\lambda, q, u, t) = (f(\lambda, q, u, t), 0)^{\text{T}}$, with the extended state $\tilde{q} = (\lambda, q)$. In other words, we add to the control system the equation $\lambda^{\Delta}(t) = 0$

(this is a standard method to derive a parametrized version of the PMP). Applying the PMP then yields an adjoint vector $\tilde{p} = (p_\lambda, p)$, where p clearly satisfies all conclusions of Theorem 1 (except (12)), and $p_\lambda^\Delta(t) = -\frac{\partial H}{\partial \lambda}(\lambda^*, q^*(t), u^*(t), p^\sigma(t), p^0, t)$ Δ -almost everywhere. From this last equation it follows that $p_\lambda(b) - p_\lambda(a) = -\int_{[a, b^*]} \frac{\partial H}{\partial \lambda}(\lambda^*, q^*(t), u^*(t), p^\sigma(t), p^0, t) \Delta t$, and then (14) follows from the already established transversality conditions.

References

- [1] R.P. Agarwal and M. Bohner. Basic calculus on time scales and some of its applications. *Results Math.*, 35(1-2):3–22, 1999.
- [2] R.P. Agarwal, M. Bohner, and A. Peterson. Inequalities on time scales: a survey. *Math. Inequal. Appl.*, 4(4):535–557, 2001.
- [3] R.P. Agarwal, V. Otero-Espinar, K. Perera, and D.R. Vivero. Basic properties of Sobolev’s spaces on time scales. *Adv. Difference Equ.*, Art. ID 38121, 14, 2006.
- [4] A.A. Agrachev and Y.L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004.
- [5] F.M. Atici, D.C. Biles, and A. Lebedinsky. An application of time scales to economics. *Math. Comput. Modelling*, 43(7-8):718–726, 2006.
- [6] Z. Bartosiewicz and D.F.M. Torres. Noether’s theorem on time scales. *J. Math. Anal. Appl.*, 342(2):1220–1226, 2008.
- [7] J. Blot and H. Chebbi. Discrete time Pontryagin principles with infinite horizon. *J. Math. Anal. Appl.*, 246(1):265–279, 2000.
- [8] M. Bohner. Calculus of variations on time scales. *Dynam. Systems Appl.*, 13(3-4):339–349, 2004.
- [9] M. Bohner and G.S. Guseinov. Double integral calculus of variations on time scales. *Comput. Math. Appl.*, 54(1):45–57, 2007.
- [10] M. Bohner and A. Peterson. *Dynamic equations on time scales. An introduction with applications*. Birkhäuser Boston Inc., Boston, MA, 2001.
- [11] M. Bohner and A. Peterson. *Advances in dynamic equations on time scales*. Birkhäuser Boston Inc., Boston, MA, 2003.
- [12] V.G. Boltyanskii. *Optimal control of discrete systems*. John Wiley & Sons, New York-Toronto, Ont., 1978.
- [13] B. Bonnard, M. Chyba, *The role of singular trajectories in control theory*. Springer Verlag, 2003.
- [14] B. Bonnard, L. Faubourg, and E. Trélat. *Mécanique céleste et contrôle des véhicules spatiaux*, volume 51 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2006.
- [15] J.M. Borwein and Q.J. Zhu. *Techniques of variational analysis*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 20. Springer-Verlag, New York, 2005.

- [16] L. Bourdin and E. Trélat. Cauchy-Lipschitz theory for shifted and non shifted Δ -Cauchy problems on time scales. *Preprint arXiv:1212.5042v1 [math.OC]*.
- [17] A. Bressan and B. Piccoli. *Introduction to the mathematical theory of control*, volume 2 of *AIMS Series on Applied Mathematics*. Springfield, MO, 2007.
- [18] J.a.e. Bryson and Y.C. Ho. *Applied optimal control*. Hemisphere Publishing Corp. Washington, D. C., 1975. Optimization, estimation, and control, Revised printing.
- [19] F. Bullo, A.D. Lewis, *Geometric control of mechanical systems. Modeling, analysis, and design for simple mechanical control systems*. Texts in Applied Mathematics, 49, Springer-Verlag, New York, 2005.
- [20] A. Cabada and D.R. Vivero. Criteria for absolute continuity on time scales. *J. Difference Equ. Appl.*, 11(11):1013–1028, 2005.
- [21] A. Cabada and D.R. Vivero. Expression of the Lebesgue Δ -integral on time scales as a usual Lebesgue integral: application to the calculus of Δ -antiderivatives. *Math. Comput. Modelling*, 43(1-2):194–207, 2006.
- [22] M.D. Canon, J.C.D. Cullum, and E. Polak. *Theory of optimal control and mathematical programming*. McGraw-Hill Book Co., New York, 1970.
- [23] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.
- [24] L.C. Evans and R.F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
- [25] L. Fan and C. Wang. *The discrete maximum principle: a study of multistage systems optimization*. John Wiley & Sons, New York, 1964.
- [26] R.A.C. Ferreira and D.F.M. Torres. Higher-order calculus of variations on time scales. In *Mathematical control theory and finance*, pages 149–159. Springer, Berlin, 2008.
- [27] J.G.P. Gamarra and R.V. Solvé. Complex discrete dynamics from simple continuous population models. *Bull. Math. Biol.*, 64:611–620, 2002.
- [28] R.V. Gamkrelidze. Discovery of the maximum principle. In *Mathematical events of the twentieth century*, pages 85–99. Springer, Berlin, 2006.
- [29] G.S. Guseinov. Integration on time scales. *J. Math. Anal. Appl.*, 285(1):107–127, 2003.
- [30] T. Haberkorn and E. Trélat. Convergence results for smooth regularizations of hybrid nonlinear optimal control problems. *SIAM J. Control Optim.*, 49(4):1498–1522, 2011.
- [31] H. Halkin. A maximum principle of the Pontryagin type for systems described by nonlinear difference equations. *SIAM J. Control*, 4:90–111, 1966.
- [32] M.R. Hestenes. *Calculus of variations and optimal control theory*. Robert E. Krieger Publishing Co. Inc., Huntington, N.Y., 1980. Corrected reprint of the 1966 original.
- [33] S. Hilger. *Ein Maßkettenkalkül mit Anwendungen auf Zentrumsmannigfaltigkeiten*. PhD thesis, Universität Würzburg, 1988.
- [34] R. Hilscher and V. Zeidan. Calculus of variations on time scales: weak local piecewise C^1_{rd} solutions with variable endpoints. *J. Math. Anal. Appl.*, 289(1):143–166, 2004.

- [35] R. Hilscher and V. Zeidan. First-order conditions for generalized variational problems over time scales. *Comput. Math. Appl.*, 62(9):3490–3503, 2011.
- [36] R. Hilscher and V. Zeidan. Weak maximum principle and accessory problem for control problems on time scales. *Nonlinear Anal.*, 70(9):3209–3226, 2009.
- [37] J.M. Holtzman. Convexity and the maximum principle for discrete systems. *IEEE Trans. Automatic Control*, AC-11:30–35, 1966.
- [38] J.M. Holtzman and H. Halkin. Discretional convexity and the maximum principle for discrete systems. *SIAM J. Control*, 4:263–275, 1966.
- [39] V. Jurdjevic, *Geometric control theory*. Cambridge Studies in Advanced Mathematics, 52, Cambridge University Press, 1997.
- [40] E. B. Lee and L. Markus, *Foundations of optimal control theory*. John Wiley, New York, 1967.
- [41] X.J. Li and J.M. Yong. *Optimal control theory for infinite-dimensional systems*. Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1995.
- [42] R.M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459–467, 1976.
- [43] B.S. Mordukhovich. *Variational analysis and generalized differentiation, I: Basic theory, II: Applications*. Volumes 330 and 331 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.
- [44] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko. The mathematical theory of optimal processes. Interscience Publishers John Wiley & Sons, Inc. New York-London, 1962.
- [45] H. Schättler, U. Ledzewicz, *Geometric optimal control, theory, methods and examples*. Interdisciplinary Applied Mathematics, Vol. 38, Springer, 2012.
- [46] S.P. Sethi and G.L. Thompson. *Optimal control theory. Applications to management science and economics*. Kluwer Academic Publishers, Boston, MA, second edition, 2000.
- [47] C. Silva and E. Trélat. Smooth regularization of bang-bang optimal control problems. *IEEE Trans. Automat. Control*, 55(11):2488–2499, 2010.
- [48] E. Trélat. *Contrôle optimal, théorie & applications*. Mathématiques Concrètes. Vuibert, Paris, 2005.
- [49] Z. Zhan, S. Chen, and W. Wei. A unified theory of maximum principle for continuous and discrete time optimal control problems. *Math. Control Relat. Fields*, 2(2):195–215, 2012.
- [50] Z. Zhan and W. Wei. On existence of optimal control governed by a class of the first-order linear dynamic systems on time scales. *Appl. Math. Comput.*, 215(6):2070–2081, 2009.